

Modelização híbrida de bioprocessos com base em métodos de engenharia de conhecimento

Maria Joana Monteiro de Carvalho Peres

Dissertação para a obtenção do grau de Doutor em Engenharia Química
pela Faculdade de Engenharia da Universidade do Porto.



Universidade do Porto
Faculdade de Engenharia
FEUP

Esta tese foi orientada pelo
Professor Doutor Sebastião José Cabral Feyo de Azevedo,
Professor Catedrático do Departamento de Engenharia Química da Faculdade de
Engenharia da Universidade do Porto
e co-orientada pelo
Professor Doutor Rui Manuel Freitas Oliveira,
Professor Auxiliar do Departamento de Química da Faculdade de Ciências e Tecnologia
da Universidade Nova de Lisboa

Abril de 2005

*Ao meu Marido, António José, e aos meus Filhos,
João Rodrigo e Diogo.*

Resumo

Esta tese identifica e desenvolve formas alternativas de modelização de bioprocessos. É abordada de forma particular a metodologia de modelização híbrida, no sentido de integrar diversas formas de conhecimento, tendo a preocupação da sistematização desta metodologia. Investigaram-se estruturas de modelos híbridos adequadas a processos biológicos e estudaram-se métodos de identificação de parâmetros.

É definida uma estrutura híbrida dinâmica geral válida para uma larga classe de problemas de modelização de biorreactores. Esta estrutura permite descrever um biorreactor por um conjunto de balanços materiais e a população celular é representada por uma mistura ajustável de representações mecanísticas e não paramétricas. São derivadas as condições de estabilidade 'entrada limitada - saída limitada' para esta estrutura híbrida por forma a garantir a positividade das concentrações em concordância com a realidade física. São comparadas duas estratégias para a identificação dos componentes não paramétricos: uma baseada no erro dos mínimos quadrados ao nível das cinéticas de reacção e outra baseada no erro dos mínimos quadrados ao nível das concentrações. Estas técnicas são ilustradas e validadas com dois casos de estudo de simulação: a produção de proteína recombinante com culturas de *Saccharomyces cerevisiae* em modo semicontínuo e o processo de produção de fermento de padeiro.

Posteriormente propõe-se uma estrutura híbrida mais complexa baseada em redes de mistura de peritos (*mixture of experts* - ME) para modelizar o 'sistema célula'. Esta rede consiste numa estrutura com dois ou mais módulos que competem entre si para formar a saída da estrutura, os quais são mediados por uma unidade de integração. A identificação dos parâmetros baseia-se no método da máxima verosimilhança, tendo sido empregue o algoritmo da Esperança-Maximização (*Expectation-Maximization* - EM). São comparados os resultados obtidos com as estruturas mais frequentes para modelizar os componentes não paramétricos, nomeadamente as redes de Percepção de Camada Múltipla (*Multiple Layer Perceptron* - MLP) e as redes de Funções de Base Radial (*Radial Basis Function* - RBF). Estes métodos foram aplicados ao processo de fermento de padeiro com dados experimentais e dados simulados e, ao processo de remoção de fósforo de águas residuais por lamas activadas com dados simulados. Demonstrou-se que as redes ME detectam a transição entre estados metabólicos distintos e que cada perito é capaz de descrever, individualmente, cada um dos estados metabólicos. Concluiu-se que as redes de mistura de peritos podem constituir um avanço na extracção de informação a partir de dados experimentais produzindo modelos mais exactos e com melhor capacidade de extrapolação no contexto da modelização híbrida.

São definidas duas estruturas híbridas baseadas em redes de mistura de peritos em que o sistema biorreactor é modelizado por balanços materiais e o 'sistema célula' é modelizado por uma representação mecanística e uma rede de mistura de peritos. A diferença entre as duas estruturas reside na ponderação de peritos, isto é, uma faz a ponderação ao nível das cinéticas de reacção, a outra faz a ponderação ao nível das concentrações. Esta última obrigou à modificação do algoritmo EM. Estes métodos foram validados com dados experimentais dum processo de produção de Polihidroxialcanoatos por culturas mistas.

É proposto um método novo de integrar a informação obtida dos diferentes (sub)modelos disponíveis acerca dum processo. Este método híbrido permite misturar peritos baseados em diferentes paradigmas de modelização e tem o mérito de obter uma combinação óptima entre os diversos modelos/fontes de conhecimento acerca do processo em estudo. Aplicando este método garante-se que em cada instante o modelo mais exacto é usado para calcular a saída final superando os métodos híbridos existentes na literatura que não entram em consideração com o verdadeiro desempenho de cada modelo nas diferentes regiões do espaço das entradas. Este método foi validado com dados experimentais dum processo de produção de fermento de padeiro.

Assim, as principais contribuições deste trabalho consistem não só no aprofundamento da base teórica da modelização híbrida como também na construção de estruturas de modelos híbridos adequados a processos biológicos, perspectivando a sua integração em metodologias avançadas de optimização e controlo de bioprocessos.

Abstract

Alternative methods of bioprocess modelling are identified and developed in this thesis. The focus is in hybrid modelling through knowledge integration having in mind the systematisation of this methodology. Hybrid modelling structures, designed for biological processes, were investigated and parameter identification methods were studied.

A general dynamic hybrid structure, valid for a wide class of problems of bioreactor modelling, is defined. This structure allows the description of the bioreactor system by a set of mass balance equations where the cell population system is represented by an adjustable mixture of non-parametric and mechanistic representations. Bounded input bounded output (BIBO) stability conditions are derived for this hybrid structure which assures the positiveness of concentrations in accordance to the physical process. Two strategies for the identification of embedded non-parametric components are compared: one based on the least square errors of kinetic reactions and another one based on the least square errors of concentrations. These techniques are illustrated and validated with two simulation case studies: the fed-batch production of recombinant protein by *Saccharomyces cerevisiae* cultures and a Baker's yeast production process.

Next a more complex hybrid structure based on mixture of experts networks (ME) is proposed for modelling the cell system. These networks consist on a structure of two or more modules, mediated by an integration unit, that compete between themselves to form the final system output. The parameter identification method follows the maximum likelihood formulation along the Expectation-Maximisation (EM) algorithm. The results obtained are compared with the most used structures for modelling the non-parametric components, such as Multiple Layer Perceptron (MLP) and the Radial Basis Functions (RBF) networks. These methods were applied to the Baker's yeast production process with simulated and experimental data, and to the simulation of wastewater phosphorus removal treatment process by activated sludge. It was demonstrated that the ME network detects the switch between metabolic pathways and each expert developed expertise in modelling each metabolic pathway. This study concluded that the mixture of experts network may represent an advance in the extraction of information from experimental data yielding more accurate models with better extrapolation properties in the context of hybrid modelling.

Two hybrid structures based on mixture of experts networks, where the bioreactor system is modelled by a set of mass balance equations and the cell system is modelled by a mechanistic term and mixture of experts network, were defined. The difference between these two structures lies in the experts weighing, i.e., one of them takes care of the weighing at the reaction kinetics level and the other one the weighing at the concentrations level. The latter demanded a modification in the EM algorithm. These methods were validated with experimental data from a mixed culture cultivation process for the production of Polyhydroxyalkanoates.

A new method of weighing the information obtained from the different available (sub-)models of the process was proposed. This method allows mixing experts based on different modelling paradigms and has the merit of searching for the optimal combination among the available models/sources of knowledge of the underlying process. Its application guarantees that at each instant the most accurate model is used to form the final output of the system, outperforming the existing methods in the literature that don't take into account the performance of each model in different regions of the input space. This method was validated with experimental data from a Baker's yeast production process.

The main contributions of this work consist not only on a better understanding of the hybrid modelling theoretical basis but also on the development of adequate hybrid model structures for biological processes, seeking its integration on advanced model-based bioreactor optimisation and control strategies.

Résumé

Cette thèse identifie et développe des formes alternatifs de modélisation de bioprocessus. La méthodologie de modélisation hybride est abordé, dans le sens d'intégration des formes de connaissance diverses, aient la préoccupation de systématiser cette méthodologie. On a recherché des structures de modèles hybrides pour des procédés biologiques et des méthodes d'identification de paramètres ont été étudiées.

On détermine une structure hybride dynamique généralisable valide pour une large classe de problèmes de modélisation de bioréacteurs. Cette structure permet de décrire un bioréacteur par un ensemble de bilans matériels et la population cellulaire est représentée par un mélange ajustable des représentations mécanistes et non-paramétriques. Les conditions de stabilité 'entrée limitée, sortie limitée' sont dérivées pour cette structure hybride, pour garantir la positivité des concentrations telle que la réalité physique. Deux stratégies sont comparées afin d'identifier les composants non-paramétriques : une basée sur l'erreur des minimums carrés au niveau des cinétiques de réaction et l'autre basée sur l'erreur des minimums carrés au niveau des concentrations. Ces techniques sont illustrées et validées avec deux cas d'étude de simulation : la production de protéine de recombinaison avec des cultures de *Saccharomyces cerevisiae* dans une manière semi-continue et le processus de production de ferment de boulanger.

Ultérieurement on propose une structure hybride plus complexe basée sur des réseaux de mélange d'experts (*mixture of experts* - ME) pour modéliser le 'système cellule'. Ces réseaux consistent en une structure avec deux ou plusieurs modules qui rivalisent entre eux pour former la sortie de la structure, qui sont négociés par une unité d'intégration. L'identification des paramètres se base sur la méthode de la vraisemblance maximale, ayant été employé l'algorithme de Espérance-Maximisation (*Expectation-Maximization* - EM). Les résultats obtenus sont comparés avec les structures les plus fréquentes pour modéliser les composantes non paramétriques, notamment les réseaux de Perception de Couche Multiple (*Multiple Layer Perception* - MLP) et les réseaux de Fonctions de Base Radiale (*Radial Basis Function* - RBF). Ces méthodes ont été appliquées au processus de ferment de boulanger avec des données expérimentales et des données simulées et, au processus de déplacement de phosphore d'eaux résiduelles par des boues activées avec des données simulées. Il s'est démontré que les réseaux ME détectent la transition entre des états métaboliques distincts et que chaque expert est capable de décrire, individuellement, chacun des états métaboliques. Il s'est conclu que les réseaux de mélange d'experts peuvent constituer une avancée dans l'extraction d'informations à partir de données expérimentales en produisant des modèles plus exacts et avec meilleure capacité d'extrapolation dans le contexte de la modélisation hybride.

Deux structures hybrides basées sur des réseaux de mélange d'experts dans lesquels le système bioréacteur est modélisé par des bilans matériels et le 'système cellule' est modélisé par une représentation mécaniste et un réseau de mélange d'experts sont définies. La différence entre les deux structures réside dans la pondération d'experts, c'est-à-dire, une fait la pondération au niveau des cinétiques de réaction, l'autre fait la pondération au niveau des concentrations. Cette dernière a obligé à la modification de l'algorithme EM. Ces méthodes ont été validées avec des données expérimentales d'un processus de production de Polihydroxyalcanoates par des cultures mixtes.

On propose une méthode innovante de pondération de l'information obtenue des différents (sous-)modèles disponibles concernant un processus. Cette méthode a été construite en se basant sur les réseaux de mélange d'experts et a le mérite d'essayer d'obtenir une combinaison optimale entre différents modèles/sources de connaissance par rapport à l'étude du processus. L'application de cette méthode assure que à chaque instant le modèle le plus exact est utilisé pour calculer la sortie finale, surmontant les méthodes existantes dans la littérature qui ne considèrent pas le véritable accomplissement de chaque modèle, dans différentes régions de l'espace des entrées. Cette méthode a été validée avec des données expérimentales d'un processus de production de ferment de boulanger.

Ainsi, les contributions principales de ce travail ne consistent seulement en approfondir la base théorique de la modélisation hybride, mais aussi à la construction de structures de modèles hybrides convenables à des processus biologiques, mettant en perspective leur intégration dans des méthodologies avancées d'optimisation et contrôle de bioprocessus.

Agradecimentos

Começo por agradecer ao meu orientador, o Professor Sebastião José Cabral Feyo de Azevedo, do Departamento de Engenharia Química da Faculdade de Engenharia da Universidade do Porto pelo tema actual que me propôs, pela confiança que em mim depositou, pelas palavras de incentivo sempre presentes, pelos conhecimentos que me transmitiu, e pelo auxílio e disponibilidade que sempre demonstrou para o fazer.

Quero também agradecer ao meu co-orientador, o Professor Rui Oliveira, do Departamento de Química da Universidade Nova de Lisboa, pelo seu profundo empenhamento na supervisão deste trabalho, e igualmente pelos conhecimentos que me transmitiu, e pelo apoio e disponibilidade que sempre demonstrou para o fazer.

Queria também expressar o meu agradecimento à Professora Doutora Ascensão Miranda Reis, Professora Auxiliar do Departamento de Química da Faculdade de Ciências e Tecnologia da Universidade Nova de Lisboa pelo facto de ter disponibilizado os dados do processo dos Bioplásticos.

Ao Director da Faculdade de Engenharia da Universidade do Porto, na pessoa do Professor Carlos Costa, devo agradecer a dispensa de serviço do CICA que me concedeu durante parte da preparação desta tese.

Devo agradecer também ao Director do Departamento de Engenharia Química da FEUP, o Professor Sebastião José Cabral Feyo de Azevedo, pelas condições de trabalho que me proporcionou, nomeadamente o espaço e o equipamento necessários à realização e escrita desta tese.

Gostaria ainda de agradecer ao CICA pelas facilidades de acesso à Internet e realçar o esforço que a Biblioteca da FEUP tem feito ao longo dos últimos anos para disponibilizar à comunidade da FEUP bibliografia de grande relevância científica e actualizada.

Também não queria deixar de agradecer à Professora Lúcia Ribeiro, enquanto Directora do CICA, pela compreensão e pela autonomia possível que de alguma forma me

concedeu durante o tempo que estive de serviço no CICA para a realização desta tese.

A ti, Cristina Faria, queria agradecer a compreensão e a força que sempre me deste para levar a bom termo esta tarefa.

Queria também agradecer aos meus amigos, colegas de gabinete, não só pela maneira carinhosa como me receberam na vossa sala, mas também pelo apoio e incentivo que sempre me demonstraram, nomeadamente à Petia, à Cristina, ao Ricardo, ao Nuno, ao Peter e ao Hélder.

À Joana Azeredo gostaria de agradecer a sua disponibilidade e amabilidade para tratar de diversos assuntos de secretariado.

Finalmente, um enorme e sincero agradecimento a todos os familiares e amigos que, sempre me souberam apoiar e transmitir palavras de conforto e incentivo.

Índice

Resumo	iii
Abstract	v
Résumé	vii
Agradecimentos	ix
Índice	xi
Lista de Figuras	xv
Lista de Tabelas	xxiii
Nomenclatura	xxv
1 Introdução	1
1.1 Modelização de bioprocessos	1
1.2 Uma nova perspectiva: modelização através da integração de conhecimento	4
1.3 Motivação	6
1.4 Objectivos e organização da tese	8
Referências	12
2 Métodos de Modelização Híbrida	17
2.1 Introdução	17
2.2 Estruturas híbridas em série	19
2.3 Estruturas híbridas em paralelo	20
2.4 Estruturas híbridas complexas	21
2.4.1 Métodos de ponderação baseados em técnicas de agrupamento	24
2.4.2 Métodos de ponderação baseados em sistemas difusos	26
2.5 Conclusões	28
Referências	29

3	Descrição de Casos de Estudo	35
3.1	Caso de Estudo I: produção de proteína recombinante em modo semi-contínuo	35
3.2	Caso de Estudo II: processo de produção de fermento de padeiro	39
3.2.1	Modelo de simulação da produção de fermento de padeiro	39
3.2.2	Dados experimentais	43
3.3	Caso de Estudo III: processo de remoção de fósforo de águas residuais por lamas activadas	44
3.4	Caso de Estudo IV: processo de produção de Polihidroxialcanoatos . . .	49
3.5	Conclusões	52
	Referências	53
4	Modelização Híbrida de Processos (Bio)químicos: Definição de Uma Estratégia	59
4.1	Introdução	60
4.2	Modelo Híbrido Geral	62
4.2.1	Redes de Percepção de Camada Múltipla (MLP)	64
4.2.2	Redes de Funções de Base Radial (RBF)	66
4.3	Derivação das condições de estabilidade BIBO	67
4.4	Identificação de Parâmetros	68
4.4.1	Estratégia I: erro dos mínimos quadrados em ρ	69
4.4.2	Estratégia II: erro dos mínimos quadrados em \mathbf{c}	70
4.5	Caso de Estudo I: produção de proteína recombinante em modo semi-contínuo	72
4.5.1	Descrição do processo	72
4.5.2	Estrutura do modelo híbrido	73
4.5.3	Resultados da identificação	74
4.6	Caso de Estudo II: processo de produção de fermento de padeiro	79
4.6.1	Descrição do processo	79
4.6.2	Estrutura do modelo híbrido	80
4.6.3	Resultados da identificação	81
4.6.4	Resultados da estabilidade BIBO	85
4.7	Conclusões	87
	Referências	88
5	Modelização de Cinéticas de Microrganismos com Mistura de Peritos	93
5.1	Introdução	94
5.2	Redes Mistura de Peritos	95

5.2.1	Arquitectura da mistura de peritos	98
5.2.2	O algoritmo da Esperança-Maximização	101
5.3	Caso de Estudo II: processo de produção de fermento de padeiro	104
5.3.1	Descrição do processo	104
5.3.2	Resultados por simulação	105
5.3.3	Resultados experimentais	107
5.3.4	Erro na vizinhança da transição	109
5.4	Caso de Estudo III: processo de remoção de fósforo de águas residuais por lamas activadas	112
5.4.1	Descrição do processo	112
5.4.2	Modelos neuronais	112
5.4.3	Comparação entre as duas estruturas de rede	114
5.5	Conclusões	118
	Referências	118
6	Modelização Híbrida Balanço Material/Mistura de Peritos	123
6.1	Introdução	124
6.2	Desenvolvimento dum modelo híbrido	126
6.2.1	Hipóteses simplificativas	126
6.2.2	Equações de balanço material	126
6.2.3	Estrutura do modelo híbrido	127
6.2.3.1	Estratégia I: minimização dos erros nas cinéticas . . .	128
6.2.3.2	Estratégia II: minimização dos erros nas concentrações	128
6.3	Resultados da identificação	130
6.3.1	Estratégia I: minimização dos erros nas cinéticas	130
6.3.2	Estratégia II: minimização dos erros nas concentrações	139
6.4	Conclusões	146
	Referências	146
7	Mistura Híbrida de Peritos Baseados em Diferentes Paradigmas de Mo- delização	149
7.1	Introdução	150
7.2	Uma nova estrutura híbrida: rede modular baseada em conhecimento .	151
7.2.1	Definição dos Peritos	152
7.2.2	Definição do Sistema de Ponderação	153
7.2.3	Identificação de parâmetros	156
7.3	Caso de Estudo II: processo de produção do fermento de padeiro	157
7.3.1	Formulação do problema	157

7.4	Resultados da identificação	161
7.5	Conclusões	163
	Referências	164
8	Conclusões e trabalho futuro	167
8.1	Estrutura de modelo híbrido	169
8.2	Identificação de parâmetros	173
8.3	Aplicabilidade aos processos biotecnológicos	174
8.4	Desenvolvimentos futuros	175
	Referências	175
	Lista de referências utilizadas na Tese	179

Lista de Figuras

1.1	Propriedades de generalização e interpolativas	2
1.2	Estrutura hierárquica do conhecimento dum processo, níveis de sofisticação dos seus componentes e resolução de detalhes (adaptado de Lubbert e Simutis, 1994).	4
1.3	Usabilidade de diferentes métodos para engenharia do conhecimento e resolução de problemas dependendo dos dados disponíveis e das teorias existentes sobre um problema (adaptado de Kasabov, 1996)	6
2.1	Estrutura híbrida em série (adaptado de Thompson e Kramer, 1994).	19
2.2	Exemplo de estrutura híbrida modular complementar para modelizar bi- orreactor agitado: ξ , concentrações de metabolitos (vector de estado) r , cinéticas de reacção, u , vector de variáveis de controlo (Adaptado de Schubert et al., 1994a).	20
2.3	Estrutura híbrida em paralelo (adaptado de Thompson e Kramer, 1994).	21
2.4	Estrutura híbrida modular competitiva+complementar (Adaptado de Si- mutis et al., 1995): F , Q , entradas conhecidas, ξ vector de estado, r vector de cinéticas. Modelo híbrido dinâmico dum processo de produ- ção de cerveja.	22
2.5	Estrutura híbrida modular competitiva+complementar (Adaptado de Si- mutis et al., 1997). Modelo para controlo óptimo de um processo de produção de penicilina.	22
2.6	Espaço bidimensional de entradas na ANN.	24
2.7	Medida de extrapolação (ε) (Medida de extrapolação $0 \leq \varepsilon \leq 1$ ob- tida por agrupamento dos padrões de treino, da Figura 2.6, com um conjunto de grupos hipersféricos (algoritmo k-média, ver detalhes em Leonard et al., 1992).	25
2.8	Estrutura híbrida competitiva que usa a medida de extrapolação ε para ponderar um modelo cinético ANN e um modelo cinético de segurança (tipo MONOD). Quando ε é grande (ANN não extrapola) a ANN é usada preferencialmente. Quando ε é baixo (a ANN está a extrapolar) o modelo de segurança é usado preferencialmente.	25

2.9	Peso relativo do modelo competitivo em função da medida de extrapolação.	28
3.1	Representação esquemática simplificada das 3 vias metabólicas para o crescimento da <i>S. cerevisiae</i> em glucose e etanol (a) metabolismo global para o crescimento em glucose (b) fermentação da glucose (c) oxidação da glucose (d) oxidação do etanol (adaptado de Oliveira, 1997)	40
3.2	Ciclo de 'fartura' e 'fome'	49
3.3	Esquema da operação do processo de produção de PHB por culturas mistas.	51
4.1	Estrutura do modelo híbrido geral para biorreactores agitados	63
4.2	Arquitectura duma rede de percepção de camada múltipla com duas camadas internas e um neurónio de saída	65
4.3	Arquitectura duma rede de funções de base radial	67
4.4	Resultados do modelo híbrido para a corrida de treino: (a) biomassa; (b) glucose; (c) proteína total; (d) proteína segregada. As linhas a cheio representam os valores 'medidos' e as linhas a tracejado representam as saídas do modelo híbrido	75
4.5	Resultados do modelo híbrido para a corrida de teste: (a) biomassa; (b) glucose; (c) proteína total; (d) proteína segregada. As linhas a cheio representam os valores 'medidos' e as linhas a tracejado representam as saídas do modelo híbrido	76
4.6	Resultados da modelização cinética: (a) taxa específica de crescimento; (b) taxa específica de produção de proteína; (c) taxa de secreção de proteína. As linhas a cheio representam as cinéticas verdadeiras e as linhas a tracejado representam as saídas do modelo híbrido	77
4.7	Cinéticas da secreção da proteína no meio: (a) taxa de secreção volumétrica; (b) taxa específica de secreção de proteína. As linhas a cheio representam as 'verdadeiras' funções e as linhas a tracejado representam as taxas do modelo. A linha ponto traço representa o h_3	78
4.8	Erro quadrático médio (MSE) em função do tempo de computação (CPU, s), obtido num processador Pentium II: para a partição de treino (-, linha a cheio) e para a partição de validação (-, linha a tracejado)	82

4.9	Resultados do modelo híbrido para o processo de produção do fermento de padeiro para os três conjuntos de dados usados para treino: (a) conjunto de treino 1; (b) conjunto de treino 2; (c) conjunto de treino 3. Os símbolos representam os valores 'medidos' e as linhas representam as saídas do modelo híbrido: +, concentração de glucose, *, concentração de etanol, o, concentração de biomassa.	83
4.10	Resultados do modelo híbrido para o processo de produção do fermento de padeiro para os três conjuntos de dados usados para validação: (a) conjunto de validação 1; (b) conjunto de validação 2; (c) conjunto de validação 3. Os símbolos representam os valores 'medidos' e as linhas representam as saídas do modelo híbrido: +, concentração de glucose, *, concentração de etanol, o, concentração de biomassa.	84
4.11	Resultados do modelo híbrido usando discretizações um passo à frente com períodos de amostragem de 0.1 (linha ponto traço), 0.25 (linha a tracejado) e 1 h (linha pontuada), e usando a formulação contínua integrada com o algoritmo de Runge-Kutta de 4 ^a /5 ^a ordem (linha a cheio), símbolo o, valores medidos.	85
4.12	Teste das condições de estabilidade BIBO do sistema (4.3) - (4.4):(a) resultados da modelização da glucose para todos os conjuntos de dados com $\mathbf{H} = \text{diag}([X \ X \ X])$ e (b) resultados da modelização da glucose para todos os conjuntos de dados com $\mathbf{H} = \text{diag}([XS \ XS \ XE])$. . .	86
5.1	Diagrama de blocos da rede mistura de peritos: as saídas dos peritos são mediadas pelo sistema de ponderação	98
5.2	Resultados para 6 corridas simuladas: (a) taxa específica de crescimento estimada com uma rede ME com 2 peritos (18 parâmetros): valores medidos (o, pontos), valores estimados (-, linha)	106
5.3	Resultados para 6 corridas simuladas: saídas da unidade de integração localizada: g_1 (... , linha pontuada), g_2 (-, linha sólida) versus concentrações de S (o, pontos brancos). A verdadeira transição dá-se para substrato constante igual a 0.0422 (g/L)	107
5.4	Resultados para 5 corridas com dados experimentais: (a) taxa específica de crescimento estimada com uma rede ME com 2 peritos (18 parâmetros): valores medidos (o, pontos), valores estimados (-, linha)	108
5.5	Resultados para 5 corridas com dados experimentais: saídas da unidade de integração localizada: g_1 (... , linha pontuada), g_2 (-, linha sólida) versus concentrações de S (•, pontos pretos). A transição detectada dá-se para substrato na vizinhança de 0.1 (g/L)	109

5.6	O quadrado do erro da estimação da taxa específica de crescimento com: (a) uma rede ME com com 2 peritos MLP (18 parâmetros);(b) uma rede MLP com 17 parâmetros	110
5.7	(a) Saídas da unidade de integração. (b) Verdadeira transição para substrato igual a 0.0422 (g/L)	111
5.8	Estrutura da rede modular para modelizar as cinéticas de consumo/produção de S_{O_2} , S_F , S_A , S_{PO_4} , X_S e X_{PAO} com dois peritos MLP ($\mathbf{S} = \{S_{O_2}, S_F, S_A, S_{PO_4}, X_S\}$, $\mathbf{q} = \{q_{SO_2}, q_{SF}, q_{SA}, q_{SPO_4}, q_{XS}, \mu\}$, nh é o número de nodos da camada escondida)	114
5.9	Erro quadrático médio obtido com uma rede MLP (o, pontos), com uma rede ME (*, asteriscos)e com uma rede ME sem incluir os parâmetros da unidade de integração (+, sinais mais): (a) partição de treino ;(b) partição de validação	115
5.10	Tempo de CPU em segundos gasto no treino numa: rede MLP (o, pontos), rede ME (*, asteriscos)	116
5.11	Saídas do sistema de ponderação: (a) partição de treino (-, linha a cheio);(b) partição de validação (-, linha a tracejado)	116
5.12	Resultados para uma partição de validação: valores medidos (o, pontos), rede MLP (-, linha sólida), rede ME (-, linha tracejada)	117
6.1	Estrutura do modelo híbrido que combina redes ME com equações de balanço material.	125
6.2	Resultados obtidos para $\rho(1)$, $\rho(2)$ e q_{PHB} para a partição de treino (caso1): valores experimentais (o, pontos), modelo ME (-, linha). . .	131
6.3	Resultados obtidos para $\rho(1)$, $\rho(2)$ e q_{PHB} para a partição de validação (caso1): valores experimentais (o, pontos), modelo ME (-, linha). . .	132
6.4	Saídas do sistema de ponderação localizado para a partição de treino (caso1): g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de NH_4 (o, pontos brancos) e X (●, pontos pretos)	132
6.5	Saídas do sistema de ponderação localizado para a partição de validação (caso1): g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de NH_4 (o, pontos brancos) e X (●, pontos pretos) . .	133
6.6	Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de treino (caso1): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).	134
6.7	Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de validação (caso1): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).	134

6.8	Resultados obtidos para $\rho(1)$, $\rho(2)$ e q_{PHB} para a partição de treino (caso2): valores experimentais (o, pontos), modelo ME (-, linha). . .	136
6.9	Resultados obtidos para $\rho(1)$, $\rho(2)$ e q_{PHB} para a partição de validação (caso2): valores experimentais (o, pontos), modelo ME (-, linha). . .	136
6.10	Saídas do sistema de ponderação localizado para a partição de treino (caso2): g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de HAc (o, pontos brancos) e PHB (●, pontos pretos) . . .	137
6.11	Saídas do sistema de ponderação localizado para a partição de validação (caso2): g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de HAc (o, pontos brancos) e PHB (●, pontos pretos)	137
6.12	Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de treino (caso2): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).	138
6.13	Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de validação (caso2): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).	138
6.14	Erro quadrático médio (MSE) em função do número de iterações (caso1): para a partição de treino (-, linha a cheio) e para a partição de validação (-, linha a tracejado).	140
6.15	Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de treino (caso1): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).	140
6.16	Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de validação (caso1): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).	141
6.17	Saídas do sistema de ponderação localizado para a partição de treino (caso1): g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de NH_4 (o, pontos brancos) e X (●, pontos pretos)	141
6.18	Saídas do sistema de ponderação localizado para a partição de validação (caso1): g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de NH_4 (o, pontos brancos) e X (●, pontos pretos) .	142
6.19	Erro quadrático médio (MSE) em função do número de iterações para a partição de treino (caso2): (-, linha a cheio) e para a partição de validação (-, linha a tracejado).	143
6.20	Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de treino (caso2): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).	144

6.21	Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de validação (caso2): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).	144
6.22	Resultados para a partição de treino (caso2): saídas do sistema de ponderação localizado: g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de HAc (o, pontos brancos) e PHB (•, pontos pretos)	145
6.23	Resultados para a partição de validação (caso2): saídas do sistema de ponderação localizado: g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de HAc (o, pontos brancos) e PHB (•, pontos pretos)	145
7.1	Diagrama de blocos da rede KBM composta por K peritos. As saídas de cada perito são mediadas por um sistema de ponderação. Os diferentes peritos expressam diferentes tipos de conhecimento: de caixa branca, de caixa cinzenta e de caixa preta	152
7.2	Sistema de ponderação: representação esquemática dum conjunto hiperesférico de grupos distribuído equidistantemente num espaço de entradas bidimensional. A região cinzenta define o subespaço T onde existem medidas do processo em estudo.	154
7.3	Sistema de ponderação: exemplo da associação entre grupos ($N_C = 20$) e peritos ($K = 2$) para o caso unidimensional do espaço das entradas. A linha a cheio representa o subconjunto $\{C\}_1$ de $N_T/2 + N_B$ grupos associados ao perito 1 ($\tau_j = 1$); a linha a tracejado representa o subconjunto $\{C\}_2$ de $N_T/2$ grupos associados ao perito 2 ($\tau_j = 2$)	154
7.4	Rede KBM para a predição da biomassa num processo de fermentação do fermento de padeiro com $K = 2$ peritos. O perito 1 é um modelo mecanístico (caixa branca) e o perito 2 é um modelo híbrido (caixa cinzenta)	158
7.5	Distribuição dos grupos no sistema de ponderação da rede KBM para o fermento de padeiro. Número total de grupos é $N_C = 55$; $N_T = 33$ grupos pertencendo ao subconjunto T e $N_B = 22$ grupos pertencendo ao subconjunto B	159

7.6	Curvas de predição da biomassa como função do tempo de cultura num processo semicontínuo de fermentação do fermento de padeiro.(a)-(e) Resultados para 5 fermentações. Os círculos representam os valores de biomassa medidos. As linhas a ponteados e tracejados representam os resultados da predição dos peritos 1 e 2, respectivamente, antes de aplicar o algoritmo EM. As linhas a cheio representam os resultados da predição pela rede KBM. (f) Saídas do sistema de ponderação em função da taxa de alimentação da glucose, g_1 (linha a tracejado) referente ao perito 1, g_2 (linha a cheio) referente ao perito 2	162
8.1	Estruturas híbridas estudadas neste trabalho.	170

Lista de Tabelas

2.1	Regras Difusas	27
3.1	Taxas específicas de crescimento do fermento de padeiro	42
3.2	Parâmetros cinéticos (tirados de Sonnleitner e Kappeli, 1986)	42
3.3	Rendimentos estequiométricos (tirados de Pomerleau e Perrier, 1990) .	43
3.4	Composição do meio de cultura para produção de fermento de padeiro .	43
3.5	Condições experimentais	44
3.6	Condições da fermentação	44
3.7	Componentes considerados no modelo simplificado	46
3.8	Parâmetros do processo	48
4.1	Número de iterações, Tempo de CPU, Erro de Validação e Erro de Treino em função dos métodos de discretização e contínuos utilizados	84
5.1	Valores iniciais das variáveis de estado do modelo simplificado	113
6.1	Erros de validação absolutos médios para as duas estratégias de identi- ficação, para os vários casos estudados	143

Nomenclatura

Caracteres Romanos Maiúsculos

A,B	matrizes
<i>B</i>	subconjunto de grupos onde não existem medidas disponíveis
<i>C</i>	conjunto de N_C grupos distribuídos equidistantemente em S_{inp}
$\{C\}_i$	subconjunto de grupos associado ao perito i
<i>C</i>	concentração de dióxido de carbono dissolvido (g/L)
<i>CTR</i>	taxa de transferência de dióxido de carbono
<i>D</i>	taxa de diluição (1/h)
<i>D_{out}</i>	espaço das saídas dos dados medidos
<i>E</i>	concentração de etanol (g/L)
$F(\mathbf{x})$	função que expressa a saída final duma rede MLP ou RBF
$F_{me}(\mathbf{x})$	função que expressa a saída final duma rede ME
<i>F</i>	taxa de alimentação de entrada (L/h)
F_S	taxa de alimentação de glucose por unidade de volume (g/(L.h))
H(c)	uma matriz de expressões cinéticas conhecidas de dimensão $r \times r$
<i>HAc</i>	concentração de acetato no processo dos bioplásticos (C-mmol/L)
I	matriz identidade
<i>J</i>	função do erro

K	número de peritos da rede de mistura de peritos
$K_{O_2}^L$	coeficiente de saturação/inibição do oxigénio (g- O_2 /m ³)
K_X	coeficiente de saturação para COD em partículas (g-COD/g-COD)
K_P	coeficiente de saturação da síntese do polifosfato (g-P/m ³)
K_A	coeficiente de saturação do acetato (g-COD/m ³)
K_{fPHA}	coeficiente de saturação do f_{PHA} (g-COD/g-COD)
K_{O_2}	coeficiente de saturação/inibição do oxigénio (g- O_2 /m ³)
K_{PO_4}	coeficiente de saturação do fósforo (g-P/m ³)
K_{NH_4}	coeficiente de saturação da amónia (g-N/m ³)
K_{PP}	coeficiente de saturação na hidrólise do polifosfato (g-P/m ³)
K_{PHA}	coeficiente de saturação do PHA (g-COD/m ³)
K_{GLY}	coeficiente de saturação do glicogénio (g-COD/m ³)
K_{fGLY}	coeficiente de saturação de f_{GLY} (g-COD/g-COD)
K_{fPP}	coeficiente de saturação de f_{PP} (g-P/g-COD)
K_{ALK}	coeficiente de saturação da alcalinidade (mol- HCO_3^- /m ³)
K_e	parâmetro de saturação para o etanol (g/L)
K_i	parâmetro de inibição (g/L)
K_s	parâmetro de saturação para a glucose (g/L)
K	matriz de coeficientes de rendimento de dimensão $n \times r$
\mathcal{L}	função de verosimilhança para o conjunto dados medidos
\mathcal{L}_c	função de verosimilhança para o conjunto completo de dados
N(c, W)	matriz de funções
NH_4	concentração de amónia no processo dos bioplásticos (N-mmol/L)
N	concentração de amónia (g/L)
N_B	número de grupos no subconjunto B

N_C	número de grupos no conjunto C
N_T	número de grupos no subconjunto T
OTR	taxa de transferência de oxigénio
O	concentração de oxigénio dissolvido (g/L)
P	número de padrões medidos
$P(\mathbf{x}, \mathbf{m}_j, \mathbf{\Sigma}_j)$	distribuição Gaussiana
$P(\mathbf{d}^t \mathbf{x}^t, \mathbf{w}_j)$	probabilidade condicional do padrão de saída \mathbf{d} dada a entrada \mathbf{x} e dado o perito j
PHB	concentração de poli- β -hidroxibutirato (C-mmol/L)
P_t	concentração total de proteína (g/L)
P_m	concentração de proteína segregada (g/L)
Q	valor esperado de \mathcal{L}_c dado \mathcal{X}
S_{O_2}	concentração de oxigénio dissolvido (g O_2 /m ³)
S_F	concentração de substrato fermentável (g COD/m ³)
S_A	concentração de acetato (g COD/m ³)
S_{NH_4}	concentração de amónia (g N/m ³)
S_{PO_4}	concentração de fosfato (g P/m ³)
S_I	concentração de fracção inerte (g COD/m ³)
S_{ALK}	alcalinidade em bicarbonato (g HCO_3^- /m ³)
S	concentração de glucose (g/L)
S_o	concentração de substrato na corrente de entrada (g/L)
S_{inp}	espaço das entradas dos dados medidos
T	subconjunto de grupos onde existem medidas disponíveis
V	volume líquido (L)
\mathbf{W}	vector de parâmetros

$\mathbf{W}_1, \mathbf{W}_2$	matrizes de parâmetros associados às ligações entre os nodos da rede neuronal MLP
$\mathbf{B}_1, \mathbf{B}_2$	matrizes de parâmetros associados às ligações entre os nodos da rede neuronal MLP
$\mathbf{W}_{1,j}$	matriz de parâmetros das ligações entre os nodos da camada 1 e 2 da rede neuronal MLP j
$\mathbf{W}_{2,j}$	matriz de parâmetros das ligações entre os nodos da camada 2 e 3 da rede neuronal MLP j
\mathcal{X}	conjunto de dados medidos
X_S	concentração de substrato lentamente biodegradável (g COD/m ³)
X_{PAO}	concentração de biomassa activa (g COD/m ³)
X_{PP}	concentração de polifosfato armazenado em PAO (g P/m ³)
X_{PHA}	concentração de PHA armazenado em PAO (g COD/m ³)
X_{GLY}	concentração de glicogénio armazenado em PAO (g COD/m ³)
X	concentração de biomassa (g/L)
\mathcal{Y}	conjunto de dados completo: dados medidos mais o conjunto de variáveis \mathcal{Z}
Y_{PHA}	coeficiente de rendimento da acumulação anaeróbia de PHA (g-COD/g-COD)
Y_{PO4}	coeficiente de rendimento $PO4/HAc$ (g-P/g-COD)
Y_{PHA}^O	coeficiente de rendimento PHA /biomassa (g-COD/g-COD)
Y_{GLY}^O	coeficiente de rendimento glicogénio/biomassa (g-COD/g-COD)
Y_{PP}^O	coeficiente de rendimento polifosfato/biomassa (g-P/g-COD)
Y_{XS}^o	coeficiente de rendimento biomassa/glucose na fase oxidativa em glucose (g biomassa /g glucose)
Y_{XS}^r	coeficiente de rendimento biomassa/glucose na fase redutiva em glucose (g biomassa /g glucose)

Y_{xe}^{oe}	coeficiente de rendimento biomassa/etanol na fase oxidativa em etanol (g biomassa /g etanol)
Y_{xe}^r	coeficiente de rendimento biomassa/etanol na fase redutiva em glucose (g biomassa /g etanol)
Y_{xo}^o	coeficiente de rendimento biomassa/oxigénio na fase oxidativa em glucose (g biomassa /g oxigénio)
Y_{xo}^{oe}	coeficiente de rendimento biomassa/oxigénio na fase oxidativa em etanol (g biomassa /g oxigénio)
Y_{xc}^o	coeficiente de rendimento biomassa/dióxido de carbono na fase oxidativa em glucose (g biomassa /g dióxidoC)
Y_{xc}^r	coeficiente de rendimento biomassa/dióxido de carbono na fase redutiva em glucose (g biomassa /g dióxidoC)
Y_{xc}^{oe}	coeficiente de rendimento biomassa/dióxido de carbono na fase oxidativa em etanol (g biomassa /g dióxidoC)
\mathcal{Z}	conjunto de variáveis omissas fictícias

Caracteres Romanos Minúsculos

\mathbf{a}_j	vector de parâmetros do sistema de ponderação associados ao perito j
$\arg \max_{\mathbf{W}} f(\mathbf{W})$	valor de \mathbf{W} para o qual a função $f(\mathbf{W})$ é máxima
$\arg \min_{\mathbf{W}} f(\mathbf{W})$	valor de \mathbf{W} para o qual a função $f(\mathbf{W})$ é mínima
$\mathbf{b}_{1,j}$	vector de parâmetros de desvio associados à camada 1 da rede neuronal MLP do perito j
$\mathbf{b}_{2,j}$	vector de parâmetros de desvio associados à camada 2 da rede neuronal MLP do perito j
\mathbf{c}	um vector de n concentrações
$\mathbf{c}_{m,t}$	valores medidos das concentrações
\mathbf{d}	vector de variáveis de saída

$diag$	diagonal duma matriz
$dim(\mathbf{W})$	dimensão do vector \mathbf{W}
exp	exponencial
\mathbf{e}_t	resíduo
f_i	função contínua não linear
f_p	taxa específica de expressão de proteína (u.a./(g biomassa.h))
f_S	fracção do substrato lentamente biodegradável (g-COD/g-COD)
f_{SI}	fracção de COD inerte produzido por hidrólise (g-COD/g-COD)
f_{PP}^{max}	fracção máxima de polifosfato (g-P/g-COD)
f_{PHA}^{max}	fracção máxima de PHA (g-COD/g-COD)
f_{GLY}^{max}	fracção máxima de glicogénio (g-COD/g-COD)
g_{PP}	factor de redução na formação de polifosfato
g_j	saída j do sistema de ponderação
k_h	coeficiente da taxa de hidrólise (g-COD/g-COD.d)
h_3	força motriz
$h_j^{p,t}$	probabilidade posterior para o perito j para o padrão t na iteração p
\mathbf{h}^p	matriz das probabilidades posteriores na iteração p
k_{PHA}	taxa de degradação de PHA (g-COD/g-COD.d)
k_{PP}	taxa de formação de polifosfato (g-P/g-COD.d)
k_{GLY}	taxa de formação de glicogénio (g-COD/g-COD.d)
\ln	logaritmo natural
m_1	número de funções de base radial
m_{AN}	taxa de manutenção anaeróbia (g-P/g-COD.d)
m_{O_2}	taxa de manutenção aeróbia (g- O_2 /g-COD.d)
\mathbf{m}_j	centros da função de distribuição Gaussiana

n_d	dimensão do espaço das saídas
n_p	número total de padrões medidos
n_x	dimensão do espaço das entradas
nh	número de nodos da camada interna
nw	número de parâmetros
q_s	taxa específica de consumo de glucose (g glucose/(g biomassa.h))
q_{omax}	taxa específica máxima de consumo de oxigénio (g O_2 /(g biomassa.h))
q_{smax}	taxa específica máxima de consumo de glucose (g glucose/(g biomassa.h))
$q_{S,AN}^{max}$	taxa máxima anaeróbia de consumo de acetato (g-COD/g-COD.d)
\mathbf{r}	vector de cinéticas
$s(x)$	função sigmóide ou tangente hiperbólica
t	tempo
$tanh$	tangente hiperbólica
\mathbf{t}_i	centros da função de base radial
\mathbf{u}	um vector de taxas volumétricas de entrada (taxas volumétricas de alimentação de nutrientes mais taxas volumétricas de transferência de massa gás-líquido) (g/(L.h))
\mathbf{x}	vector de entradas
\mathbf{y}	estimativa/predição dada pelo modelo global
\mathbf{y}_{ANN}	estimativa/predição dada por uma rede ANN
$\mathbf{y}_{segurança}$	estimativa/predição dada pelo modelo de segurança
\mathbf{y}_j	elemento j do vector \mathbf{y}

Caracteres Gregos Maiúsculos

Γ	matriz diagonal de ganhos do estimador
Λ_i	matriz de covariâncias do perito i associada a θ_i
Σ	matriz das covariâncias
Φ	taxa de excreção de proteína (1/h)
Ω	matriz diagonal de ganhos do estimador

Caracteres Gregos Minúsculos

α_j	parâmetros escalares
ε	medida de extrapolação da rede neuronal
η_{fe}^L	factor de redução da hidrólise anaeróbia
θ_i	centro do grupo mais próximo da entrada \mathbf{x} pertencente ao subconjunto $\{C\}_i$
θ	conjunto de parâmetros total da rede ME
μ	taxa específica de crescimento (1/h)
μ_{exp}	taxa específica de crescimento 'medida' (1/h)
μ_s^o	taxa específica de crescimento respirativo em glucose (1/h)
μ_s^r	taxa específica de crescimento fermentativo em glucose (1/h)
μ_e^o	taxa específica de crescimento respirativo em etanol (1/h)
μ_{emax}	taxa específica máxima de crescimento oxidativo em etanol (1/h)
ν	grupos definidos como funções Gaussianas
$\rho(\mathbf{c})$	um vector de r funções cinéticas desconhecidas
σ	desvio padrão
τ	vector de parâmetros inteiros

v_{pm}	taxa de secreção volumétrica (u.a./((L.h))
$v(\mathbf{c})$	um vector de taxas volumétricas de reacção (g/(L.h))
φ	função de base radial

Índices e Expoentes

in	corrente de entrada
out	corrente de saída
a	denota a partição com r variáveis de estado
max	denota valor máximo
\wedge	significa quantidade estimada
T	transposto
p	iteração p

Abreviaturas¹

ANN	redes neuronais artificiais (<i>Artificial Neural Network</i>)
ATCC	<i>American Type Culture Collection</i>
ATP	Adenosina Tri-fosfato (<i>Adenosine-Tri-Phosphate</i>)
ASM	Modelo das Lamas Activadas (<i>Activated Sludge Model</i>)
BIBO	estabilidade entrada limitada saída limitada (<i>Bounded Input Bounded Output</i>)
CART	Árvores de Regressão e Classificação (<i>Classification and Regression Trees</i>)
CG	Gradiente Conjugado (<i>conjugate gradient</i>)

¹Os acrónimos utilizados são baseados em acrónimos de língua inglesa de uso corrente em Portugal pois facilita a leitura e a compreensão deste texto.

COD	Carência Química de Oxigênio (<i>Chemical Oxygen Demand</i>)
CPU	unidade de processamento central (<i>central processing unit</i>)
DA	método do recozimento determinístico (<i>Deterministic Annealing</i>)
EM	Esperança-Maximização (<i>Expectation-Maximization</i>)
ER	retículo endoplasmático
HME	mistura hierárquica de peritos (<i>Hierarchical mixtures of experts</i>)
KBH	híbrido baseado em conhecimento (<i>Knowledge Based Hybrid</i>)
KBM	rede modular baseada em conhecimento (<i>Knowledge Based Modular network</i>)
ME	mistura de peritos (<i>mixture of experts</i>)
MLP	perceptrão de camada múltipla (<i>Multiple Layer Perceptron</i>)
MLR	regressão múltipla linear (<i>Multiple Linear Regression</i>)
MSE	erro quadrático médio (<i>Mean square error</i>)
NADH	forma reduzida do dinucleotido da nicotinamida adenina (<i>Nicotinamide Adenine Dinucleotide Hydrogen</i>)
NARMAX	média móvel autoregressiva não linear com entrada exógena (<i>Non-linear autoregressive moving average with exogenous input</i>)
NLPCA	análise de componentes principais não-linear (<i>Non-linear principal component analysis</i>)
PHA	Polihidroxialcanoatos
PHB	Poli- β -hidroxibutirato
OBE	estimador baseado num observador (<i>observer-based estimator</i>)
ODE	equação diferencial ordinária (<i>ordinary differential equation</i>)
PAO	microrganismos acumuladores de fósforo (<i>Phosphorous Accumulating Organisms</i>)
PCR	regressão de componentes principais (<i>Principal Component Regression</i>)

RBF	funções de base radial (<i>Radial Basis Function</i>)
rDNA	Ácido Desoxirribonucleico recombinado
RK	<i>Runge-Kutta</i>
rpm	rotações por minuto
SBR	reactor descontínuo sequencial (<i>Sequencing Batch Reactor</i>)
slpm	<i>standard liter per minute</i>
SOM	<i>self-organizing map</i>
SQP	programação quadrática sequencial (<i>Sequential Quadratic Programming</i>)
TCA	ciclo dos ácidos tricarboxílicos
u.a.	unidades arbitrárias

Capítulo 1

Introdução

Conteúdo do Capítulo

Esta tese identifica e desenvolve formas alternativas de modelização de bioprocessos. É abordada de forma particular a metodologia de modelização híbrida no sentido de integrar diversas formas de conhecimento. Neste capítulo são detalhados os motivos que orientaram e levaram à prossecução do desenvolvimento desta tese. Os objectivos específicos são discriminados e por fim, é apresentada a estrutura da tese com um pequeno resumo do conteúdo de cada capítulo.

1.1 Modelização de bioprocessos

Embora podendo ser mais ou menos formal, qualquer modelo têm na sua génese o propósito de concentrar num padrão um conjunto de observações (Ljung, 1987). Assim, a construção e utilização de modelos matemáticos baseados em dados observados é desde há muito aceite como metodologia científica básica. Daí que, com o progresso da tecnologia digital, a modelização computacional e as aplicações baseadas em modelos têm-se desenvolvido progressivamente, sendo actualmente reconhecidas como áreas de grande prioridade para o futuro (Edgar, 1996). As questões que se colocam são: que modelos e que aplicações?

A abordagem clássica da engenharia química (e bioquímica) favorece os modelos baseados nos chamados 'primeiros princípios' que, como afirma Villiermaux (1996), são os mesmos de há cem anos. Actualmente, muito pelas dificuldades experimentadas na

análise dos processos bioquímicos (e porque há meios técnicos para conduzir estudos noutras direcções) é universalmente aceite que o saber quantitativo mecanístico é por vezes insuficiente, podendo ser compensado ou complementado com outras formas de conhecimento - estatístico, qualitativo, difuso ou eminentemente heurístico.

A abordagem convencional na modelização de processos baseia-se em equações de balanço de massa, energia e, se necessário, momento e população. Esta forma de modelização requer ainda conhecimento de leis cinéticas, bem como de propriedades termodinâmicas, de transporte e físicas.

Frequentemente, a capacidade preditiva dos modelos clássicos dos processos biológicos é bastante limitada. Tal deve-se às características intrinsecamente não lineares e variantes no tempo do metabolismo celular, com estruturas cinéticas que no melhor dos casos apenas em parte são conhecidas, mas que frequentemente, são mesmo desconhecidas de todo.

Independentemente da inquestionável relevância dos estudos em curso, com base em modelos mecanísticos complexos, novos métodos baseados em técnicas de modelização orientada para dados, estão a ser desenvolvidos. Estes são capazes de extrair conhecimento dum dado processo a partir de dados experimentais, de forma a ultrapassar as dificuldades existentes para expressar o conhecimento através de equações matemáticas com bases mecanísticas.

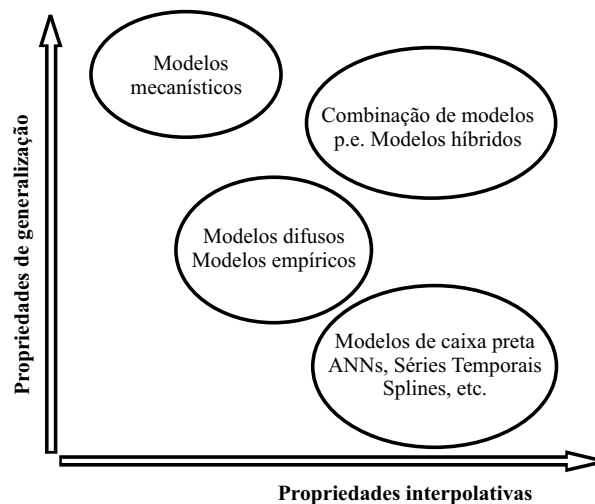


Figura 1.1: Propriedades de generalização e interpolativas

O rápido crescimento dos recursos computacionais levou ao desenvolvimento de um largo número de métodos de modelização baseados em dados. Estes métodos abrangem técnicas estatísticas bem estabelecidas, que vão desde a regressão múltipla linear (*Multiple Linear Regression* - MLR) e regressão de componentes principais (*Principal*

Component Regression - PCR), até técnicas não lineares como a análise de componentes principais não linear (*Non-linear principal component analysis* - NLPCA) e média móvel autoregressiva não linear com entrada exógena (*Non-linear autoregressive moving average with exogenous input* - NARMAX).

Nos últimos anos as redes neuronais artificiais (*Artificial Neural Network* - ANN) introduziram uma nova atitude na modelização em engenharia bioquímica (Scott e Harmon Ray, 1993; Ye et al., 1994; Montague e Morris, 1994). As ANN são, em muitos casos, capazes de representar relações multivariável - em particular as que ocorrem em sistemas dinâmicos altamente não lineares - sem qualquer conhecimento do processo subjacente. Os parâmetros da rede são estimados por treino da rede com um conhecimento apriorístico do processo.

Contudo, esta ferramenta matemática pode conduzir a previsões que violam restrições fundamentais ditadas pelos princípios de conservação, sobretudo quando opera fora dos domínios em que foi treinada (ver Figura 1.1), daí que, tenha surgido naturalmente uma nova forma de modelizar processos que tenta ultrapassar as desvantagens das duas abordagens apresentadas, combinando-as numa formulação híbrida (Psychogios e Ungar, 1992; Thompson e Kramer, 1994; Lubbert e Simutis, 1994).

A modelização híbrida procura a inclusão de todo o conhecimento disponível do processo. Os fundamentos dos modelos híbridos são os princípios da conservação. Os aspectos menos conhecidos ou desconhecidos de um processo, como as cinéticas da reacção, são modelizados com o apoio de metodologias de inteligência artificial, incluindo o conhecimento apriorístico do processo. Assim, as partes complexas e pouco conhecidas dum processo em vez de serem modelizadas pela aplicação de leis de conservação podem ser modelizadas recorrendo à utilização de ANNs. Isto é geralmente designado por modelização híbrida baseada em conhecimento (*Knowledge-based hybrid* - KBH) e está bem documentado na literatura (Roubos et al., 1999; Russell e Bakker, 1997; Schubert et al., 1994b,a; Feyer de Azevedo et al., 1997; Zorzetto et al., 2000; Peres et al., 2001).

Uma estratégia eficiente de optimização da operação de processos requer a cooperação entre todos os factores que produzem conhecimento, i.e., cientistas, operadores de processo e tecnologia. Daí que, quanto mais conhecimento houver sobre um determinado processo, mais exactos serão os modelos e mais eficientes serão as estratégias de operação baseadas em modelos.

Uma estratégia de modelização híbrida que relacione todos os factores que produzem conhecimento, e que permita a optimização do processo suportado por todos os níveis do conhecimento disponíveis, apresenta-se pois como uma técnica com elevado potencial

para aplicação na indústria.

1.2 Uma nova perspectiva: modelização através da integração de conhecimento

A modelização através da integração de conhecimento tem como objectivo explorar todas as fontes de conhecimento/informação apriorísticas acerca dum processo, conhecimento esse que deve ser incorporado no modelo do processo (Schubert et al., 1994b; Psichogios e Ungar, 1992; Thompson e Kramer, 1994; Feyerherm et al., 1997; Simutis et al., 1997).

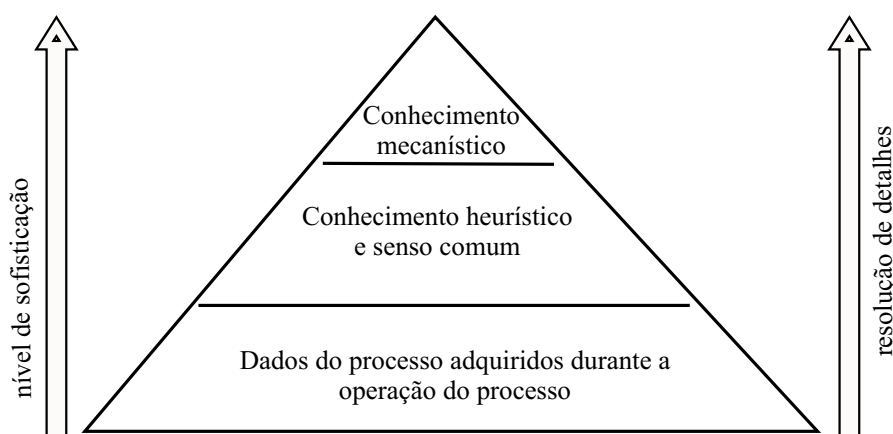


Figura 1.2: Estrutura hierárquica do conhecimento dum processo, níveis de sofisticação dos seus componentes e resolução de detalhes (adaptado de Lubbert e Simutis, 1994).

Normalmente, existe uma multiplicidade de fontes de informação em processos biotecnológicos. A estrutura hierárquica do conhecimento dum processo estende-se deste o nível mecanístico até à informação escondida nos registos dos dados do processo (Figura 1.2). Concretamente, três grandes tipos de conhecimento podem ser identificados:

1. Conhecimento mecanístico (fenomenológico): este tipo de conhecimento é duma maneira geral representado por modelos matemáticos. Esta é a abordagem clássica seguida pelos engenheiros químicos e bioquímicos para desenvolver os seus modelos de processos. Tem o nível mais elevado de sofisticação envolvendo a compreensão dos mecanismos básicos de transporte e da cinética da reacção. Estes mecanismos básicos são mal percebidos ou mesmo completamente desconhecidos, daí que,

duma maneira geral, este tipo de conhecimento é o que existe em menor quantidade.

2. Conhecimento heurístico e senso comum: este tipo de conhecimento é de natureza qualitativa existindo normalmente em grandes quantidades na indústria. A teoria de sistemas difusos é um método possível para manipular este tipo de informação pois fornece métodos que permitem quantificar o conhecimento qualitativo. O conhecimento heurístico é muitas vezes formulado por regras práticas de procedimentos. Estas podem ser representadas por sistemas de inferência difusos baseados na lógica difusa e sistemas periciais (Sugeno, 1985; Kosko, 1992; Wang, 1994). Esta forma de conhecimento deve ser igualmente considerada como um recurso importante para a modelização e controlo de processos (Hitzmann et al., 1992; Sterbacek e Votruba, 1993).
3. Dados medidos 'brutos': em muitas situações o conhecimento mecanístico e/ou heurístico não é suficiente para construir um modelo dum processo com a exactidão necessária. Nestas situações, a modelização orientada para dados pode ser utilizada para melhorar a exactidão do modelo. Em muitos processos industriais os mecanismos relevantes de causa/efeito têm sido registados ao longo dos anos na forma de dados entrada/saída. Estes arquivos de dados podem ser um recurso importante na modelização das partes desconhecidas do processo através de métodos de caixa preta, nomeadamente, através de técnicas baseadas em séries temporais e em redes neuronais artificiais (ANNs). Sjoberg et al. (1995) fez uma revisão sobre a modelização do tipo caixa preta em identificação de sistemas. Em particular, as ANNs têm suscitado um grande interesse pela parte dos investigadores nos últimos anos. Estas redes provaram ser muito flexíveis na representação de relações não lineares complexas (Cybenko, 1989; Hornik et al., 1989; Poggio e Girosi, 1990) sem necessitarem de qualquer tipo de conhecimento sobre a estrutura do modelo subjacente. Alguns resultados importantes foram publicados acerca da aplicação de ANNs a problemas de identificação e controlo de sistemas dinâmicos (Hunt et al., 1992; Pollard et al., 1992; Narendra e Parthasarathy, 1990).

Portanto, os vários tipos de conhecimento acerca da operação do processo existem e estão bem identificados. Falta saber de que maneira se podem integrar por forma a construir um modelo mais exacto do processo.

Segundo (Kasabov, 1996), na perspectiva da engenharia do conhecimento podem ser utilizados métodos diferentes para obter uma solução. A Figura 1.3 representa os vários métodos e a relação entre eles quando se utilizam na resolução de problemas. Kasabov

(1996) tece algumas considerações sobre em que casos podem ou devem ser usados cada um dos métodos:

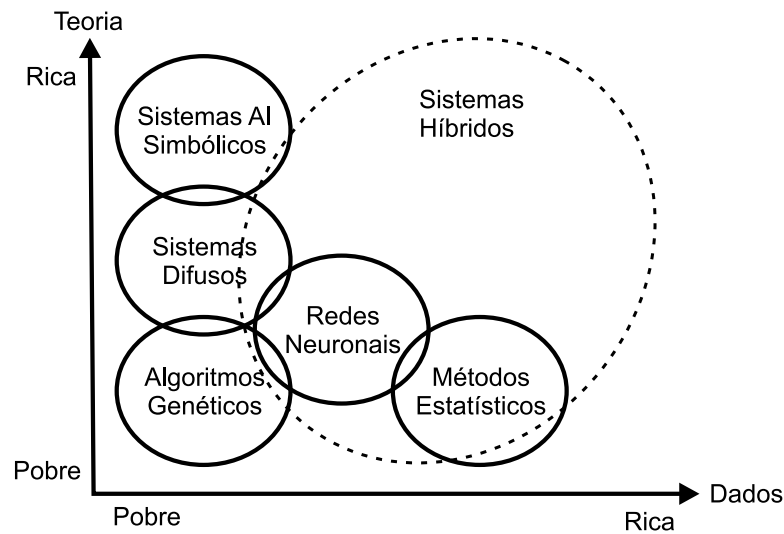


Figura 1.3: Usabilidade de diferentes métodos para engenharia do conhecimento e resolução de problemas dependendo dos dados disponíveis e das teorias existentes sobre um problema (adaptado de Kasabov, 1996)

- os métodos estatísticos podem ser utilizados quando existem dados estatisticamente representativos do problema e se conhece a função que se pretende modelizar;
- os métodos baseados em sistemas AI simbólicos podem ser utilizados quando o conhecimento do problema está definido por regras rígidas bem definidas;
- os sistemas difusos aplicam-se quando o conhecimento do problema inclui regras heurísticas;
- as redes neurais podem-se aplicar quando existem dados do problema mas não se conhece a função que se pretende modelizar;
- os algoritmos genéticos não requerem nem dados nem regras heurísticas mas apenas um critério de seleção por onde começar;
- os sistemas híbridos podem usar os diferentes métodos.

1.3 Motivação

A integração de conhecimento e de métodos baseados em modelos são ainda, até certo ponto, conceitos novos na história dos processos industriais. Todavia, estes métodos são a base de metodologias avançadas de monitorização, de optimização e de

controlo de processos químicos e bioquímicos. Apesar de tudo, constata-se que os processos industriais estão num estágio de absorção destas metodologias baseadas em modelos, em parte, como consequência do desenvolvimento da instrumentação digital e computacional.

Neste contexto, vale a pena analisar a situação nas indústrias bioquímicas com algum detalhe. Nestas indústrias vários factores contribuem para rácios de benefício/custo que não favorecem o investimento em operação de bioprocessos baseada em modelos clássicos (Royce, 1993; Simutis et al., 1993). Frequentemente, variáveis chave do processo não são medidas, por problemas de esterilização, pela inexistência de técnicas de medição fiáveis ou pelo elevado custo associado a instrumentação mais complexa. Também a complexidade dos mecanismos de crescimento dos microrganismos, de formação de produto e a complexidade do meio da reacção complicam o desenvolvimento de modelos matemáticos fiáveis. Um desafio importante de modelização será precisamente o desenvolvimento de estimativas fiáveis em linha de variáveis intracelulares em sistemas biológicos (Sonnleitner, 1999).

Num ambiente onde o conhecimento disponibilizado pelos modelos mecanísticos é escasso e os tempos de desenvolvimento são sucessivamente encurtados por razões económicas tornam-se necessárias metodologias alternativas para a modelização de processos químicos e bioquímicos. Existe pois uma motivação inequívoca para encontrar novas direcções para a modelização deste tipo de processos.

Concretamente, a modelização pode ser vista como um exercício de expressão e representação do conhecimento numa forma compacta. E porque, duma maneira geral existem diferentes tipos de conhecimento sobre o mesmo processo duas abordagens conceptuais norteiam o princípio da modelização:

1. seleccionar uma das fontes de conhecimento existentes e adoptar e explorar a correspondente técnica de modelização; ou
2. expressar e integrar todas as fontes de conhecimento existentes num só modelo híbrido do processo.

Na perspectiva da engenharia do conhecimento a segunda alternativa é mais vantajosa, uma vez que a exactidão do modelo depende da quantidade e da qualidade do conhecimento disponível. Schubert et al. (1994b,a), Feyerherm et al. (1997) e Simutis et al. (1997) mostraram que quando o conhecimento dos princípios básicos (fenomenológicos) carece de modelos matemáticos, estes podem ser complementados com

outras técnicas de modelização como, por exemplo, ANNs e/ou sistemas difusos/periciais que incorporam fontes de conhecimento não utilizadas usualmente.

Vários autores propuseram estruturas de modelização híbrida para processos químicos e bioquímicos. Concretamente, Psychogios e Ungar (1992) e Thompson e Kramer (1994) propuseram modelos híbridos que combinam modelos matemáticos com ANNs para modelizar reactores químicos e bioquímicos. Por seu turno, Simutis et al. (1993) sugeriu a combinação de ANNs com sistemas difusos. Já em Glassey et al. (1997) é sugerido um modelo que combina ANNs com sistemas periciais para supervisionar um bioprocessos. Em qualquer dos casos, as aplicações referidas são orientadas a um problema específico. Constata-se que a modelização híbrida não é um problema com uma estrutura bem definida, no sentido que ainda carece de uma teoria subjacente. Daí a necessidade de definir uma estrutura de modelo híbrido genérica que permita sistematizar o seu estudo e a sua aplicação.

1.4 Objectivos e organização da tese

O principal objectivo desta tese é o desenvolvimento de técnicas de modelização híbrida do tipo caixa cinzenta, particularmente adequadas para processos biotecnológicos tendo a preocupação da sistematização destas técnicas tornando-as passíveis de ser utilizadas independentemente do processo em estudo.

Os tópicos a investigar serão os seguintes:

1. caracterização teórica de estruturas híbridas do tipo caixa cinzenta;
2. identificação de parâmetros;
3. aplicabilidade a processos biotecnológicos.

O primeiro objectivo consiste na caracterização duma estrutura híbrida genérica que combina modelos mecanísticos com modelos não paramétricos. Propõe-se estudar a identificação e a estabilidade deste tipo de estruturas.

Normalmente os modelos não paramétricos utilizados para descrever as cinéticas são redes neurais artificiais. Contudo, dada a especificidade de determinados processos biológicos, caracterizada por terem vias metabólicas distintas, surge a necessidade de procurar alternativas para a descrição das cinéticas deste tipo de processos biológicos. Concretamente, os métodos não paramétricos alternativos desenvolvidos no âmbito deste

trabalho serão baseados em redes de mistura de peritos (*mixture of experts* - ME) (Jacobs et al., 1991).

Dadas as características intrínsecas numa rede de mistura de peritos, que serão detalhadas ao longo do trabalho, perspectiva-se que este tipo de redes será a ferramenta que faltava para integrar de forma sistemática os diferentes tipos de conhecimento existentes acerca dum processo. Daí que, este trabalho, também pretenda dar o seu contributo na definição duma mistura híbrida de peritos baseados em diferentes paradigmas de modelização. Desta forma, procura-se uma alternativa para integrar o conhecimento existente acerca dum processo mas que tem a mais valia de ser um método que não é orientado a um processo específico.

Também, será analisada a aplicabilidade dos métodos propostos baseados em estruturas híbridas a processos biológicos, nomeadamente, ao processo de produção de fermento de padeiro, a um processo de produção de proteína recombinante com a cultura de *Saccharomyces cerevisiae*, ao processo de remoção de fósforo de águas residuais por lamas activadas e ao processo de produção de Polihidroxialcanoatos por culturas mistas (plástico biodegradável sintetizado biologicamente que tem propriedades similares ao polipropileno).

Esta tese está organizada em 8 capítulos da seguinte forma:

Capítulo 1 - Introdução

Capítulo 2 - Métodos de Modelização Híbrida

Capítulo 3 - Descrição de Casos de Estudo

Capítulo 4 - Modelização Híbrida de Processos (Bio)químicos: Definição de uma Estratégia

Capítulo 5 - Modelização de Cinéticas de Microrganismos com Mistura de Peritos

Capítulo 6 - Modelização Híbrida Balanço Material/Mistura de Peritos

Capítulo 7 - Mistura Híbrida de Peritos Baseados em Diferentes Paradigmas de Modelização

Capítulo 8 - Conclusões e Trabalho Futuro

No Capítulo 2 - Métodos de Modelização Híbrida - revêm-se as principais técnicas de modelização híbrida propostas na literatura. São referidas as estruturas em série, em

paralelo e estruturas mais complexas que podem ser simultaneamente em paralelo e em série.

No Capítulo 3 - Descrição de Casos de Estudo - faz-se a descrição dos processos biológicos em estudo, a saber: processo de produção de proteína recombinante com culturas de *Saccharomyces cerevisiae* em modo semicontínuo, processo de produção de fermento de padeiro, processo de remoção de fósforo de águas residuais por lamas activadas e processo de produção de bioplásticos por culturas mistas. São também descritos os modelos matemáticos de espaço de estados adoptados neste trabalho para estes processos.

No Capítulo 4 - Modelização Híbrida de Processos (Bio)químicos: Definição de uma Estratégia - é dada uma panorâmica geral dos modelos híbridos, nomeadamente, sobre a estrutura mais utilizada e melhor estudada até ao momento. É proposta uma estrutura híbrida genérica e é feita a sua caracterização teórica. Esta estrutura híbrida do tipo caixa cinzenta combina modelos mecanísticos com modelos não paramétricos: o sistema biorreactor é descrito por um conjunto de equações de balanço material e o sistema célula é representado por uma mistura flexível de representações não paramétricas e mecanísticas. Definem-se duas estratégias de identificação de parâmetros e derivam-se as condições de estabilidade entrada limitada saída limitada (BIBO) para o modelo híbrido geral. Esta técnica é ilustrada em dois casos de estudo de simulação: processo de produção de proteína recombinante com culturas de *Saccharomyces cerevisiae* em modo semicontínuo e ao processo de produção de fermento de padeiro.

O trabalho deste capítulo deu origem à publicação Oliveira et al. (2005):

Oliveira, R., Peres, J. e Feyo de Azevedo, S. (2005), 'Hybrid modelling of fermentation processes using artificial neural networks: A study on identification and stability', em M. Pons e J. F. M. van Impe (Editores), *Computer Applications in Biotechnology 2004*, Elsevier (ISBN: 0-08-044251-X), Páginas 195 - 200.

No Capítulo 5 - Modelização de Cinéticas de Microrganismos com Mistura de Peritos - é proposta uma estrutura híbrida baseada em redes de mistura de peritos para modelizar as cinéticas de processos biológicos. São apresentadas as razões para utilizar este tipo de estruturas para modelizar processos biológicos complexos onde vários mecanismos metabólicos podem ocorrer simultaneamente. A identificação dos parâmetros é baseada na máxima verosimilhança e o algoritmo de treino é o Esperança-Maximização (EM). O modelo mistura de peritos é comparado com os modelos baseados em redes de Percepção de Camada múltipla e redes de Funções de Base Radial. Esta técnica é aplicada a dois casos de estudo: ao processo de produção de fermento de padeiro com dados simulados

e dados experimentais; ao processo de remoção de fósforo de águas residuais por lamas activadas com dados simulados onde é utilizado o modelo das Lamas Activadas 2d (Activated Sludge Model 2d, ASM2d, Henze et al., 1999).

O trabalho deste capítulo deu origem às publicações Peres et al. (2005a) e Peres et al. (2003) e ao manuscrito que se encontra em preparação para publicação Peres et al. (2005b):

Peres, J., Oliveira, R. e Foyo de Azevedo, S. (2005a), 'Hybrid modelling of fermentation processes: A study on the use of modular neural networks for modelling cells reaction kinetics', em M. Pons e J. F. M. van Impe (Editores), *Computer Applications in Biotechnology 2004*, Elsevier (ISBN: 0-08-044251-X), Páginas 293 - 298.

Peres, J., Oliveira, R. e de Azevedo, S. F. (2003), 'Modelling cells reaction kinetics with artificial neural networks: A comparison of three network architectures', em A. Kraslawski e I. Turunen (Editores), *European Symposium On Computer Aided Process Engineering - 13*, Elsevier Science Bv, volume 14 de *Computer-Aided Chemical Engineering*, Páginas 839 - 844

Peres, J., Oliveira, R. e Foyo de Azevedo, S. (2005b), 'A study on the application of modular neural networks for modelling cell reaction kinetics', em *preparação para publicação*.

No Capítulo 6 - Modelização Híbrida Balanço Material/Mistura de Peritos - são propostas duas estruturas híbridas que combinam balanços materiais com redes de mistura de peritos. Uma primeira estrutura faz a mistura de peritos ao nível das cinéticas do sistema célula e a segunda estrutura faz a mistura ao nível das concentrações do sistema biorreactor. Relativamente à primeira estrutura híbrida proposta, optou-se pela primeira estratégia descrita no Capítulo 4 para identificação de parâmetros: primeiro as cinéticas são estimadas por diferenciação numérica das curvas de concentrações e por resolução das equações de balanço material. Com os dados assim obtidos treinou-se a rede de mistura de peritos com os algoritmos usuais descritos no Capítulo 5. Relativamente à segunda estrutura híbrida proposta, optou-se pela estratégia II descrita no Capítulo 4 para a identificação dos parâmetros o que obrigou a modificar o algoritmo EM. Ambos os modelos foram validados pelo método da validação cruzada. Estas metodologias são comparadas ao modelizar um processo de produção de bioplásticos por culturas mistas à escala laboratorial.

O trabalho deste capítulo deu origem à publicação Peres et al. (2004):

Peres, J., Oliveira, R., Seram, L. S., Lemos, P., Reis, M. A. e de Azevedo, S. F.

(2004), 'Hybrid modelling of a pha production process using modular neural networks', em A. Barbosa-Póvoa e H. Matos (Editores), *European Symposium On Computer-Aided Process Engineering - 14*, Elsevier Science Bv, volume 18 de *Computer-Aided Chemical Engineering*, Páginas 733 - 738.

No Capítulo 7 - Mistura Híbrida de Peritos Baseados em Diferentes Paradigmas de Modelização - é proposta uma nova estrutura híbrida que permite integrar os diferentes tipos de conhecimento, usualmente disponíveis para a modelização de processos químicos e bioquímicos: conhecimento mecanístico, conhecimento heurístico e conhecimento escondido em registos de dados do processo. Esta estrutura é baseada em mistura de peritos e toma em consideração a exactidão de cada perito para construir a saída final do modelo. O conceito de Rede Modular Baseada em Conhecimento (Knowledge Based Modular network - rede KBM) é apresentado. O algoritmo da Esperança-Maximização (EM) é empregue para combinar de forma óptima os peritos dentro da estrutura de rede KBM. Os conceitos são ilustrados com a aplicação da produção de fermento de padeiro.

O trabalho deste capítulo deu origem à publicação Peres et al. (2001):

Peres, J., Oliveira, R. e de Azevedo, S. F. (2001), 'Knowledge based modular networks for process modelling and control', *Computers & Chemical Engineering*, 25(4-6), 783 - 791.

No Capítulo 8 - Conclusões e Trabalho Futuro - são resumidas as conclusões do trabalho e apresentadas as perspectivas futuras de trabalho.

Referências

- Cybenko, G. (1989), 'Approximation by superpositions of a sigmoidal function', *Mathematics of Control, Signals, and Systems*, **2**, 303–314.
- Edgar, T. F. (1996), 'Modelling and control - back to the future, part i', *CAST Communications*, **19**(1), 7–12.
- Feyo de Azevedo, S., Dahm, B. e Oliveira, F. R. (1997), 'Hybrid modelling of biochemical processes: A comparison with the conventional approach', *Computers & Chemical Engineering*, **21**, S751–S756.
- Glassey, J., Ignova, M., Ward, A. C., Montague, G. A. e Morris, A. J. (1997), 'Bioprocess supervision: Neural networks and knowledge based systems', *Journal of Biotechnology*, **52**(3), 201–205.
- Henze, M., Gujer, W., Mino, T., Matsuo, T., Wentzel, M. C., Marais, G. V. R. e Van Loosdrecht, M. C. M. (1999), 'Activated sludge model no.2d, asm2d', *Water Science and Technology*, **39**(1), 165–182.
- Hitzmann, B., Lubbert, A. e Schugerl, K. (1992), 'An expert system approach for the control of a bioprocess .1. knowledge representation and processing', *Biotechnology and Bioengineering*, **39**(1), 33–43.
- Hornik, K., Stinchcombe, M. e White, H. (1989), 'Multilayer feedforward networks are universal approximators', *Neural Networks*, **2**(5), 359–366.
- Hunt, K. J., Sbarbaro, D., Zbikowski, R. e Gawthrop, P. J. (1992), 'Neural networks for control-systems: a survey', *Automatica*, **28**(6), 1083–1112.
- Jacobs, R. A., Jordan, M. I., Nowlan, S. J. e Hinton, G. E. (1991), 'Adaptive mixtures of local experts', *Neural Computation*, **3**, 79–87.
- Kasabov, N. K. (1996), *Foundations of neural Networks, Fuzzy Systems, and Knowledge Engineering*, Bradford Books, The MIT Press.

- Kosko, B. (1992), *Neural networks and fuzzy systems: a dynamical system approach to machine intelligence*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Ljung, J. (1987), *System Identification - Theory for the User*, Prentice-Hall.
- Lubbert, A. e Simutis, R. (1994), 'Using measurement data in bioprocess modeling and control', *Trends in Biotechnology*, **12**(8), 304–311.
- Montague, G. e Morris, J. (1994), 'Neural-network contributions in biotechnology', *Trends in Biotechnology*, **12**(8), 312–324.
- Narendra, K. e Parthasarathy, K. (1990), 'Identification and control of dynamical systems using neural networks', *IEEE Transactions on Neural Networks*, **1**(1), 4–27.
- Oliveira, R., Peres, J. e Feyo de Azevedo, S. (2005), 'Hybrid modelling of fermentation processes using artificial neural networks: A study on identification and stability', em M. Pons e J. F. M. van Impe (Editores), *Computer Applications in Biotechnology 2004*, Elsevier (ISBN: 0-08-044251-X), Páginas 195 – 200.
- Peres, J., Oliveira, R. e de Azevedo, S. F. (2001), 'Knowledge based modular networks for process modelling and control', *Computers & Chemical Engineering*, **25**(4-6), 783–791.
- Peres, J., Oliveira, R. e de Azevedo, S. F. (2003), 'Modelling cells reaction kinetics with artificial neural networks: A comparison of three network architectures', em A. Kraslawski e I. Turunen (Editores), *European Symposium On Computer Aided Process Engineering - 13*, Elsevier Science Bv, volume 14 de *Computer-Aided Chemical Engineering*, Páginas 839–844.
- Peres, J., Oliveira, R. e Feyo de Azevedo, S. (2005a), 'Hybrid modelling of fermentation processes: A study on the use of modular neural networks for modelling cells reaction kinetics', em M. Pons e J. F. M. van Impe (Editores), *Computer Applications in Biotechnology 2004*, Elsevier (ISBN: 0-08-044251-X), Páginas 293 – 298.
- Peres, J., Oliveira, R. e Feyo de Azevedo, S. (2005b), 'A study on the application of modular neural networks for modelling cell reaction kinetics', *em preparação para publicação*.
- Peres, J., Oliveira, R., Serafim, L. S., Lemos, P., Reis, M. A. e de Azevedo, S. F. (2004), 'Hybrid modelling of a pha production process using modular neural networks', em A. Barbosa-Póvoa e H. Matos (Editores), *European Symposium On Computer-Aided Process Engineering - 14*, Elsevier Science Bv, volume 18 de *Computer-Aided Chemical Engineering*, Páginas 733–738.

- Poggio, T. e Girosi, F. (1990), 'Networks for approximation and learning', *Proceedings of the IEEE*, **78**(9), 1481–1497.
- Pollard, J. F., Broussard, M. R., Garrison, D. B. e San, K. Y. (1992), 'Process identification using neural networks', *Computers & Chemical Engineering*, **16**(4), 253–270.
- Psichogios, D. C. e Ungar, L. H. (1992), 'A hybrid neural network-1st principles approach to process modeling', *AIChE Journal*, **38**(10), 1499–1511.
- Roubos, J. A., Krabben, P., Setness, M., Babuska, R., Heijnen, J. e Verbrugen, H. B. (1999), 'Hybrid model development for fed-batch bioprocesses combining physical equations with the metabolic network and black-box kinetics', em *6th Workshop on fuzzy systems*, Brunel University, Uxbridge, Páginas 231–239.
- Royce, P. N. (1993), 'A discussion of recent developments in fermentation monitoring and control from a practical perspective', *Critical Reviews in Biotechnology*, **13**(2), 117–149.
- Russell, N. T. e Bakker, H. H. C. (1997), 'Modular modelling of an evaporator for long-range prediction', *Artificial Intelligence in Engineering*, **11**(4), 347–355.
- Schubert, J., Simutis, R., Dors, M., Havlik, I. e Lubbert, A. (1994a), 'Bioprocess optimization and control application of hybrid modeling', *Journal of Biotechnology*, **35**(1), 51–68.
- Schubert, J., Simutis, R., Dors, M., Havlik, I. e Lubbert, A. (1994b), 'Hybrid modeling of yeast production processes combination of a-priori knowledge on different levels of sophistication', *Chemical Engineering & Technology*, **17**(1), 10–20.
- Scott, G. M. e Harmon Ray, W. (1993), 'Creating efficient nonlinear neural network process models that allow model interpretation', *Journal of Process Control*, **3**(3), 163–178.
- Simutis, R., Havlik, I. e Lubbert, A. (1993), 'Fuzzy-aided neural network for real-time state estimation and process prediction in the alcohol formation step of production-scale beer brewing', *Journal of Biotechnology*, **27**(2), 203–215.
- Simutis, R., Oliveira, R., Manikowski, M., de Azevedo, S. F. e Lubbert, A. (1997), 'How to increase the performance of models for process optimization and control', *Journal of Biotechnology*, **59**(1-2), 73–89.

- Sjoberg, J., Zhang, Q. H., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P. Y., Hjalmarsson, H. e Juditsky, A. (1995), 'Nonlinear black-box modeling in system identification: A unified overview', *Automatica*, **31**(12), 1691–1724.
- Sonnleitner, B. (1999), 'Bioanalysis and biosensors for bioprocess monitoring', *Advances in Biochemical Engineering/Biotechnology* 66, (volume editor).
- Sterbacek, Z. e Votruba, J. (1993), 'An expert system applied to the control of an industrial-scale bioreactor', *Chemical Engineering Journal and The Biochemical Engineering Journal*, **51**(2), B35–B42.
- Sugeno, M. (1985), *Industrial applications of fuzzy control*, North-Holland, Amsterdam.
- Thompson, M. L. e Kramer, M. A. (1994), 'Modeling chemical processes using prior knowledge and neural networks', *AIChE Journal*, **40**(8), 1328–1340.
- Villiermaux, J. (1996), 'Future prospects for chemical engineering research and technology', *Chem. Tech. Europe*, 21–23.
- Wang, L. X. (1994), *Adaptive fuzzy systems and control: design and stability analysis*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Ye, K., Fujioka, K. e Shimizu, K. (1994), 'Efficient control of fed-batch baker's yeast cultivation based on neural network', *Process Control and Quality*, **5**(4), 245–250.
- Zorzetto, L. F. M., Maciel, R. e Wolf-Maciel, M. R. (2000), 'Process modelling development through artificial neural networks and hybrid models', *Computers & Chemical Engineering*, **24**(2-7), 1355–1360.

Capítulo 2

Métodos de Modelização Híbrida

Conteúdo do Capítulo

Neste capítulo revêem-se as principais técnicas de modelização híbrida de processos químicos e/ou bioquímicos. A estrutura de modelo híbrido mais estudada combina fenómenos de transporte (com base em balanços materiais e/ou energéticos) com técnicas de modelização não paramétricas tais como as redes neuronais artificiais. No que diz respeito à sua estrutura, os modelos híbridos podem ser essencialmente classificados como 'Modelos híbridos em série' e 'Modelos híbridos em paralelo'. Vários autores propuseram estruturas mais complexas que são simultaneamente estruturas em série e em paralelo. Neste capítulo revêem-se estas metodologias de modelização híbrida. É a partir de uma análise centrada nestas técnicas que se seleccionam, em capítulos seguintes, as técnicas mais adequadas para modelizar bioprocessos.

2.1 Introdução

A modelização matemática constitui uma ferramenta fundamental em ciências de engenharia permitindo a compreensão dos mecanismos fenomenológicos dos processos. Os modelos são ainda fundamentais como ferramentas para tomar decisões sobre a operação de processos, para controlo e optimização de processos e constituem uma ferramenta importante na passagem do desenvolvimento à escala laboratorial para a escala industrial.

Os modelos matemáticos podem ser classificados de diversas formas sob o ponto de vista da estrutura, do sistema alvo, do tipo de conhecimento que incorporam, da escala (atômica, macroscópica), da natureza estatística, se é dinâmico ou estático, se é discreto ou contínuo, se é linear ou não linear, etc. No contexto deste trabalho de doutoramento importa classificar os modelos sob o ponto de vista do conhecimento incorporado. Uma classificação adequada poderá ser: modelos de tipo caixa branca, de tipo caixa preta e de tipo caixa cinzenta em função do tipo de informação a partir do qual são construídos.

Os modelos de tipo caixa branca têm uma estrutura baseada em princípios fundamentais, cuja aplicação se traduz, no caso de processos químicos e bioquímicos, em equações de balanço material, de energia, de momento e de população, em leis cinéticas, em leis termodinâmicas que exprimem o equilíbrio químico, etc. ou em transformações matemáticas conhecidas. Portanto, são modelos representados por equações matemáticas que traduzem os mecanismos físico-químicos do comportamento do processo. Os modelos de tipo caixa branca são classificados como modelos paramétricos.

Os modelos de tipo caixa preta baseiam-se unicamente nos dados do processo. Concretamente, os modelos de caixa preta descrevem os processos através de mapeamentos de entrada/saída sem qualquer significado físico sobre o processo. Estes métodos usam métodos estatísticos para a partir de dados extrair informação acerca do processo. Redes neuronais artificiais (ANN), séries temporais, splines, regressão múltipla linear (*Multiple Linear Regression* - MLR), regressão de componentes principais (*Principal Component Regression* - PCR), a análise de componentes principais não linear (*Non-linear principal component analysis* - NLPCA) e a média móvel autoregressiva não linear com entrada exógena (*Non-linear autoregressive moving average with exogenous input* - NARMAX) são alguns exemplos de modelos de caixa preta. Este tipo de modelos são classificados como não paramétricos e, em regra, tem que possuir uma base estatística sólida.

Métodos e soluções que usam modelos de caixa preta são correntemente aceites como uma alternativa exequível ou como uma aproximação complementar para o objectivo último da representação do funcionamento do processo. Bhat e Mcavoy (1990) mostraram que a natureza não linear e variável no tempo dos processos químicos pode ser modelizada de modo exacto com tempo de desenvolvimento reduzido usando ANNs, desde que estejam disponíveis dados medidos, em quantidade e qualidade suficientes, abrangendo toda a região de operação do processo.

Os modelos de tipo caixa cinzenta resultam da combinação de modelos de caixa branca com modelos de caixa preta. A modelização híbrida é um modelo de caixa cinzenta pois integra vários tipos de conhecimento. Thompson e Kramer (1994) classificaram

este tipo de modelos como semiparamétricos. Os modelos com base qualitativa como os modelos difusos ou os sistemas periciais são também classificados como modelos do tipo caixa cinzenta.

Os métodos de modelização híbrida mais utilizados e melhor estudados combinam submodelos mecanísticos conhecidos (expressos por modelos paramétricos) com submodelos não paramétricos para descrever as partes desconhecidas do processo. Thompson e Kramer (1994) classificaram estes modelos, do ponto de vista de estrutura, em modelos híbridos em série e em paralelo. Estas duas estruturas constituem a base para a construção de modelos híbridos modulares. As próximas secções revêem estas duas estruturas, assim como outras mais complexas descritas na literatura.

2.2 Estruturas híbridas em série

A abordagem seguida na literatura para desenvolvimento de modelos híbridos tem sido uma abordagem 'modular'. O processo alvo é normalmente dividido em vários subsistemas mais simples. Numa segunda fase, são identificadas as várias formas de conhecimento disponíveis para cada subsistema. Finalmente propõe-se diagramas de blocos, nos quais os subsistemas representados por diversas formas de conhecimento, trocam informação.

A estrutura híbrida mais simples e mais amplamente utilizada em aplicações de bioprocessos é composta por dois módulos que se complementam. Esta estrutura, designada por 'Modelo híbrido em série' por Thompson e Kramer (1994) ou 'Estrutura modular complementar' em Oliveira et al. (2000) e Feye de Azevedo et al. (2001) pode ser representada genericamente da seguinte forma (Figura 2.1):

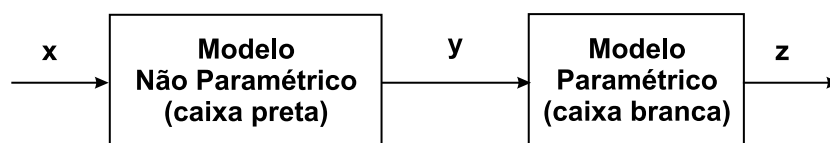


Figura 2.1: Estrutura híbrida em série (adaptado de Thompson e Kramer, 1994).

Esta estrutura adequa-se, pois, a processos sobre os quais recai conhecimento mecanístico parcial. O modelo não paramétrico serve, pois, para descrever as partes do processo 'desconhecidas' do ponto de vista mecanístico.

Esta estrutura tem sido amplamente utilizada para modelização de bioprocessos (Psychogios e Ungar, 1992; Thompson e Kramer, 1994; Montague e Morris, 1994; Feye de

Azevedo et al., 1997; van Can et al., 1998, 1999; Braake et al., 1998; Chen et al., 2000; Anderson et al., 2000; Babuska et al., 1999; Roubos et al., 1999; Karama et al., 2001a,b; Georgieva et al., 2003; Lauret et al., 2000; Molga, 2003).

Este tipo de processos quando envolve biorreactores pode naturalmente ser dividido em dois subsistemas: o subsistema macroscópico 'biorreactor' e o subsistema 'população celular'. Sobre o primeiro existe conhecimento sólido sobre os fenómenos de transporte que descrevem o subsistema. No caso das células, o conhecimento mecanístico é muito limitado, pelo que se opta por uma técnica de modelização não paramétrica (ver Figura 2.2).

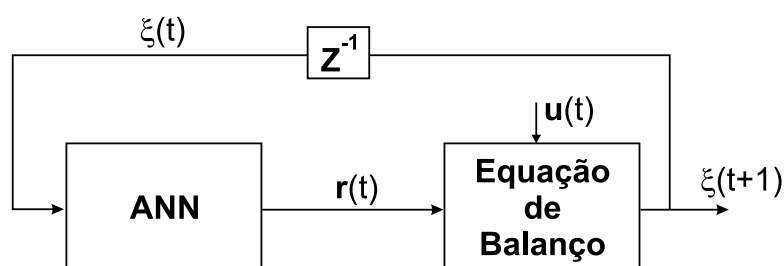


Figura 2.2: Exemplo de estrutura híbrida modular complementar para modelizar biorreactor agitado: ξ , concentrações de metabolitos (vector de estado) r , cinéticas de reacção, u , vector de variáveis de controlo (Adaptado de Schubert et al., 1994a).

2.3 Estruturas híbridas em paralelo

Na abordagem semiparamétrica paralela, as saídas da rede neuronal artificial e do modelo paramétrico são combinadas de forma a determinar a saída final do modelo (ver Figura 2.3). Este tipo de estruturas é normalmente aplicado quando se tem um modelo de caixa branca completo do processo, mas, no entanto, o modelo não tem exactidão suficiente para optimização ou controlo. A rede neuronal artificial é treinada sobre os resíduos obtidos entre os dados do processo e o modelo paramétrico para de alguma forma compensar a inexactidão do modelo mecanístico dada a complexidade intrínseca do processo.

Este tipo de estruturas tem sido também utilizado por alguns autores, nomeadamente em processos de tratamento de águas residuais (Zhao et al., 1997; Lee et al., 2002, 2005).

Uma variante à estrutura paralela, é a estrutura competitiva (Oliveira et al., 2000; Feyo de Azevedo et al., 2001). Nas estruturas competitivas existem diferentes tipos de

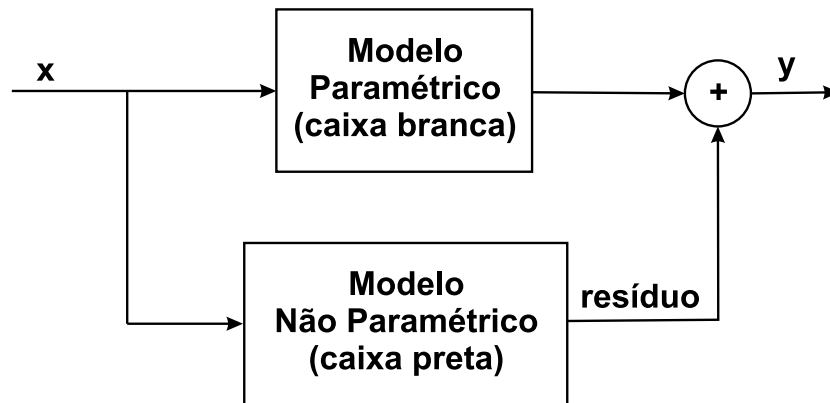


Figura 2.3: Estrutura híbrida em paralelo (adaptado de Thompson e Kramer, 1994).

conhecimento acerca do mesmo subsistema do processo, os quais competem entre si. Este é o caso quando a informação se sobrepõe. Como refere Schubert et al. (1994a) diferentes tipos de informação acerca do mesmo fenómeno devem ser usadas em simultâneo, sendo um desperdício de conhecimento desenvolver um modelo de um processo baseado só numa fonte de conhecimento das várias disponíveis acerca do subsistema em estudo. Este tópico é abordado com mais detalhe na secção seguinte.

2.4 Estruturas híbridas complexas

A estratégia modular de desenvolvimento de modelos híbridos, quando aplicado a processos complexos, poderá resultar em estruturas híbridas mais complexas que combinam os tipos de estruturas referidos nas secções anteriores, obtendo-se uma estruturas modulares híbridas que são simultaneamente competitivas e complementares. As Figuras 2.4 e 2.5 ilustram dois exemplos deste tipo de estruturas.

Outros exemplos de estruturas complexas aplicadas a bioprocessos são descritas em Senger e Karim (2003) e Eikens e Karim (1999). Nomeadamente, Senger e Karim (2003) definiram uma estrutura híbrida complementar complexa que envolve cinco redes MLP e uma função MONOD interligadas em série. Esta estrutura funciona como um sensor por programação e foi aplicada a um processo de produção de proteína recombinante.

Por sua vez, Eikens e Karim (1999) propuseram estruturas baseadas só em redes neuronais. Neste caso não há estruturas em série/paralelo, no entanto, este tipo de estruturas não são estruturas puramente de tipo caixa preta porque existe um conhecimento geral da estrutura do sistema. Concretamente, utilizaram diferentes modelos baseados em MLPs e Redes Recorrentes para modelizar os diferentes estados metabó-

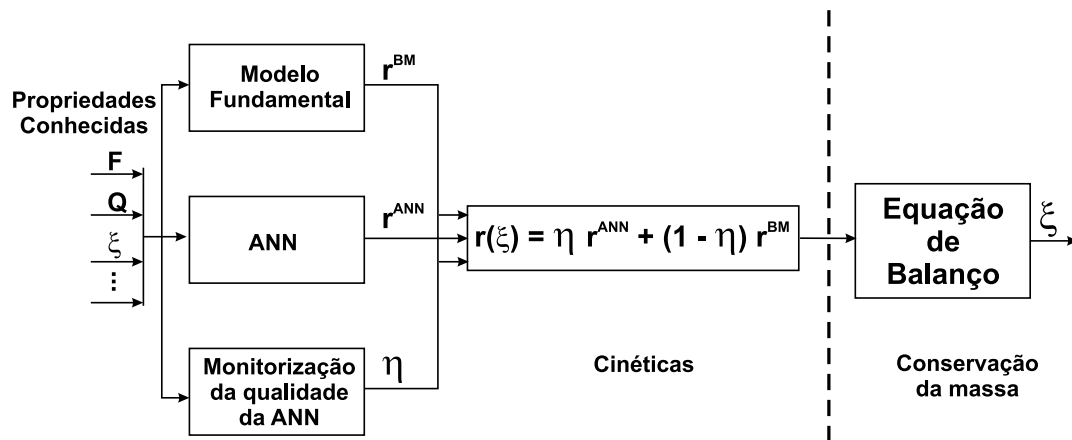


Figura 2.4: Estrutura híbrida modular competitiva+complementar (Adaptado de Simutis et al., 1995): F , Q , entradas conhecidas, ξ vector de estado, r vector de cinéticas. Modelo híbrido dinâmico dum processo de produção de cerveja.

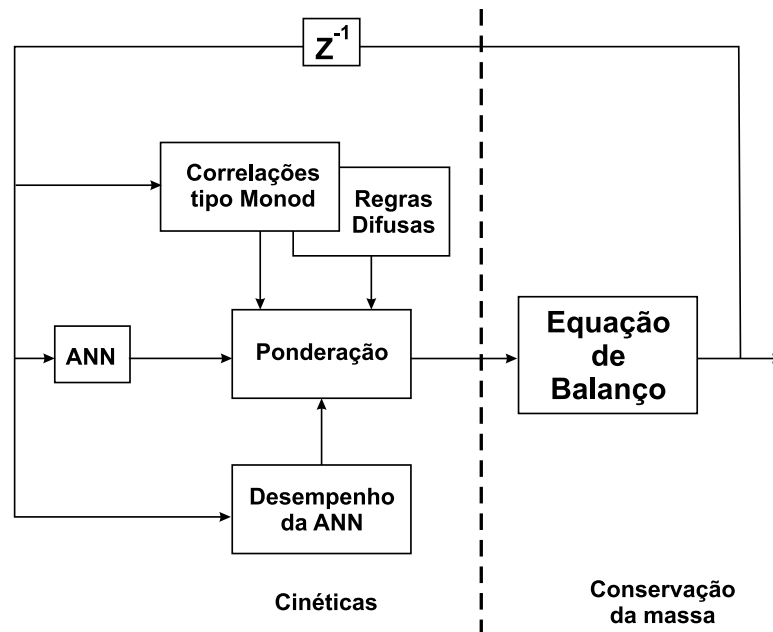


Figura 2.5: Estrutura híbrida modular competitiva+complementar (Adaptado de Simutis et al., 1997). Modelo para controlo óptimo de um processo de produção de penicilina.

licos do processo de fermento de padeiro. A decisão de usar determinado modelo em determinada região do espaço das entradas é definida pelos seguintes métodos:

1. método baseado em conhecimento a priori do processo a ser identificado.

Neste caso o espaço das entradas é dividido através de funções de interpolação suaves ou abruptas. A função de interpolação suave utilizada foi anteriormente definida em (Johansen e Foss, 1997) e é baseada em funções gaussianas. A função de interpolação abrupta é definida por limiares baseados no conhecimento do processo.

2. método baseado em técnicas não supervisionadas.

Neste caso o método implementado é similar ao proposto em Chinrungrueng (1993). Usam-se dois algoritmos baseados em técnicas de agrupamento nomeadamente, o algoritmo k-média adaptativo (Chinrungrueng e Sequin, 1995) e o SOM (*self-organizing map*) (Kohonen, 1995) para dividir o espaço das entradas em grupos. Posteriormente cada grupo é identificado por uma rede neuronal artificial.

3. método baseado em *Non-linear gated experts* (Weigend et al., 1995).

Neste caso a partição do espaço das entradas é efectuada por uma rede neuronal artificial.

A identificação baseada nos dois primeiros métodos é resolvida sequencialmente, primeiro são determinadas as partições do espaço das entradas e depois é feita a identificação de cada modelo associado a cada estado metabólico. O terceiro método é o único que permite simultaneamente particionar o espaço das entradas e identificar o modelo de cada estado metabólico. Eikens e Karim (1999) concluíram que as estimativas obtidas pelo terceiro método eram as mais exactas.

Enquanto que as estruturas complementares têm sido referidas frequentemente na literatura, as estruturas competitivas complexas não são tão comuns. Com efeito, a carência de resultados teóricos sólidos nesta área dificulta a aplicação deste tipo de estruturas para optimização e controlo de processos.

Nas estruturas competitivas, é necessário um mecanismo para ponderar cada sub-modelo. Uma questão central é a do método de ponderação. A ponderação deveria obedecer ao critério: dado um conjunto de entradas o modelo mais exacto deverá ter o peso superior no resultado final e o modelo menos exacto deverá ter o peso inferior.

Nas próximas secções são descritos os métodos de ponderação mais referidos na literatura. Nesta tese, estudar-se-ão métodos de ponderação alternativos.

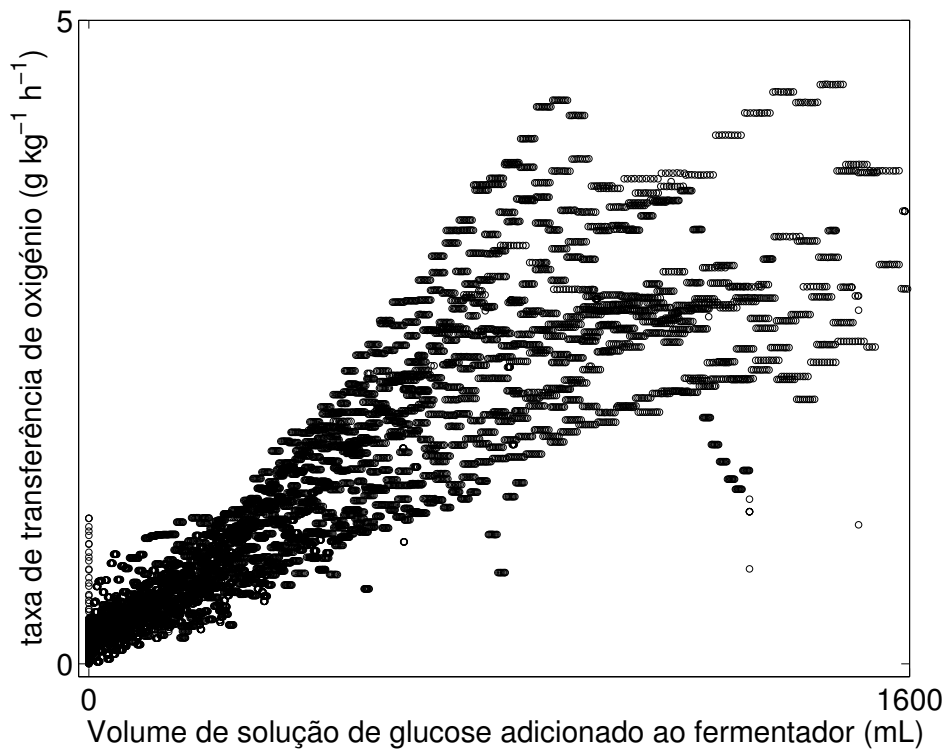


Figura 2.6: Espaço bidimensional de entradas na ANN.

2.4.1 Métodos de ponderação baseados em técnicas de agrupamento

Leonard et al. (1992) sugeriu o uso de técnicas de agrupamento (*clustering*) para monitorizar a fiabilidade das redes neuronais artificiais. Simutis et al. (1995) desenvolveu este conceito e sugeriu a aplicação de técnicas de agrupamento para combinar redes neuronais artificiais com um modelo de segurança do tipo caixa branca. O método consiste em aplicar um algoritmo de agrupamento para transformar um conjunto de medidas discretas (isto é, o espaço de entradas medidas da rede neuronal artificial designado por domínio de experiência) numa função de densidade contínua. Posteriormente, esta função é usada para avaliar a medida de extrapolação da rede neuronal - ε . Na Figura 2.6 mostra-se um exemplo de espaço de entrada na ANN e respectivas medidas de volume de solução de glucose adicionada ao fermentador e taxas de transferência de oxigénio. Na Figura 2.7 mostra-se o valor da medida de extrapolação, ε , em todo o espaço bidimensional abrangido pelas variáveis da Figura 2.6. Note-se que a zona onde a ANN tem mais influência é precisamente na zona central o que coincide com a existência de pontos medidos.

É com base neste valor ε que a decisão é tomada de ou usar o modelo de rede neuronal

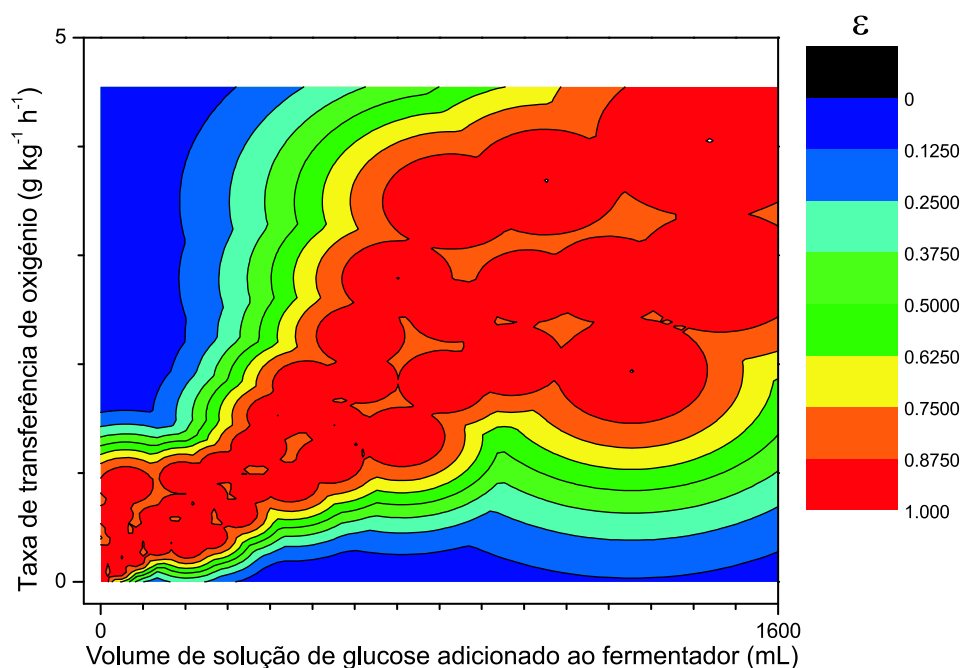


Figura 2.7: Medida de extrapolação (ϵ) (Medida de extrapolação $0 \leq \epsilon \leq 1$ obtida por agrupamento dos padrões de treino, da Figura 2.6, com um conjunto de grupos hiperesféricos (algoritmo k-média, ver detalhes em Leonard et al., 1992).

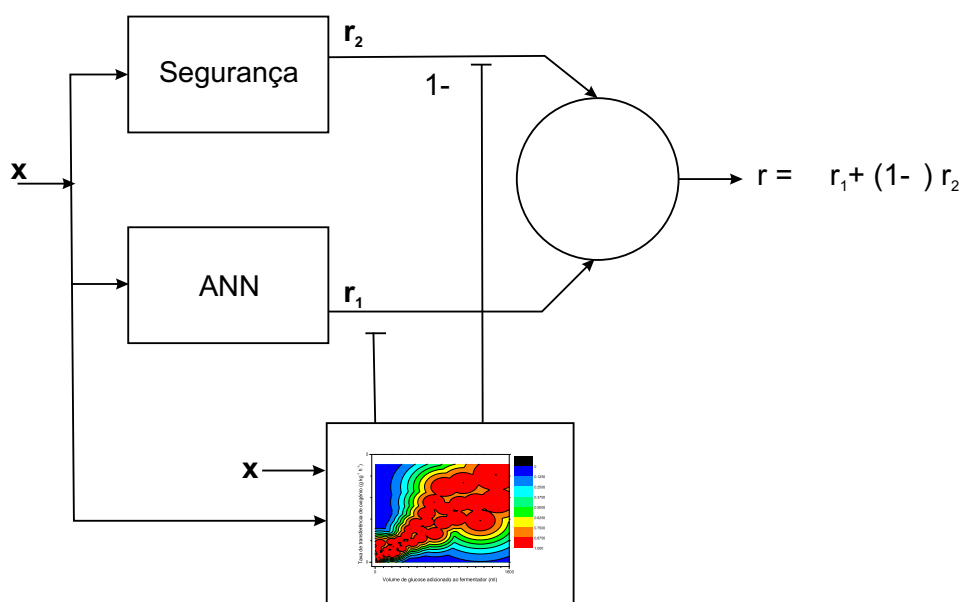


Figura 2.8: Estrutura híbrida competitiva que usa a medida de extrapolação ϵ para ponderar um modelo cinético ANN e um modelo cinético de segurança (tipo MONOD). Quando ϵ é grande (ANN não extrapola) a ANN é usada preferencialmente. Quando ϵ é baixo (a ANN está a extrapolar) o modelo de segurança é usado preferencialmente.

ou usar um modelo de segurança competitivo com melhores propriedades de extrapolação (Figura 2.8). À partida, assume-se que dentro do domínio da experiência da rede neuronal as estimativas/predições da mesma são mais exactas que qualquer outro modelo em competição e fora do domínio de experiência considera-se que as estimativas/predições não são fiáveis. Daí que, neste último caso, o modelo de segurança deve ter maior prioridade.

Concretizando matematicamente, o sistema de ponderação é baseado no valor ε da seguinte forma:

$$\mathbf{y} = \varepsilon \mathbf{y}_{ANN} + (1 - \varepsilon) \mathbf{y}_{segurança} \quad (2.1)$$

Esta forma de ponderação apresenta duas desvantagens importantes:

1. É restritiva, isto é, só pode ser usada numa estrutura com dois modelos competitivos. Daí que, uma estrutura definida com mais de dois modelos requer a aplicação de outro método;
2. O desempenho de cada modelo não é tido em conta no sistema de ponderação. Este método é baseado na presunção que a rede neuronal artificial é sempre melhor no seu domínio de experiência e pouco fiável fora deste domínio. Isto é, assume-se à partida que fora do domínio de experiência o modelo competitivo de segurança tem melhores capacidades extrapolativas e portanto deve ser preferido.

2.4.2 Métodos de ponderação baseados em sistemas difusos

Schubert et al. (1994a,b) propuseram uma estratégia de ponderação baseada em sistemas difusos e aplicaram-na ao processo de fermento de padeiro. O método é direccionado para resolver um problema específico, por isso a decisão sobre que modelo deve ser utilizado nas diferentes regiões do espaço das entradas requer conhecimento heurístico acerca do processo de ponderação. O método é baseado na avaliação, neste caso heurística, acerca das capacidades extrapolativas de cada modelo e não tem em linha de conta o verdadeiro desempenho de cada modelo em diferentes regiões do espaço das entradas. Simutis et al. (1993) já tinha aplicado esta ideia ao processo de produção de cerveja à escala industrial para descrever as diferentes fases do processo por diferentes modelos. A decisão de escolher um modelo ou outro é baseada em sistemas difusos que suavizam a transição entre as fases do processo. Também Horiuchi e Hiraga (1999) aplicou esta ideia à produção industrial de vitamina B2.

Oliveira (1998) propôs um método de ponderação que associa a medida de extrapolação ε com um sistema difuso. O método foi usado para modelizar as taxas de consumo do precursor e de amónia num processo de produção de penicilina.

Este método de ponderação consiste num sistema inferencial difuso com uma entrada - a medida de extrapolação - e três saídas correspondendo aos pesos relativos dos três modelos competitivos que descrevem as cinéticas do processo de produção de penicilina:

- rede neuronal artificial (ANN)
- modelo estequiométrico (STOI)
- correlações empíricas (CORR).

A medida de extrapolação foi atribuída com três conjuntos difusos: LOW, MED e HIGH. Cada variável de saída foi atribuída com dois conjuntos difusos: LOW e HIGH. Três regras difusas (ver Tabela 2.1) definem o mapeamento entre a medida de extrapolação e os três pesos relativos. Foi utilizada uma função de pertença baseada numa função radial:

$$\varphi = \exp\left(-\frac{(\varepsilon - \mu)^2}{\sigma^2}\right) \quad (2.2)$$

Tabela 2.1: Regras Difusas

Regra	ε	W_{ANN}	W_{CORR}	W_{STOI}
1	LOW	LOW	LOW	HIGH
2	MED	LOW	HIGH	LOW
3	HIGH	HIGH	LOW	LOW

A simulação do sistema em função da medida de extrapolação ε é apresentada na Figura 2.9. Com este conjunto de regras difusas o mecanismo de ponderação funciona da seguinte forma:

1. dentro do domínio de experiência do modelo - conjunto de pontos medidos utilizado para identificar os parâmetros do modelo - o modelo cinético ANN tem um peso relativo de $W_{ANN} = 1$,
2. fora do domínio de experiência o modelo estequiométrico mecanístico tem um peso relativo de $W_{STOI} = 1$,

3. na zona de transição entre a não extrapolação e a extrapolação, o modelo empírico da correlação tem um peso relativo de $W_{CORR} = 1$,
4. o sistema define uma transição suave entre os três modelos competitivos.

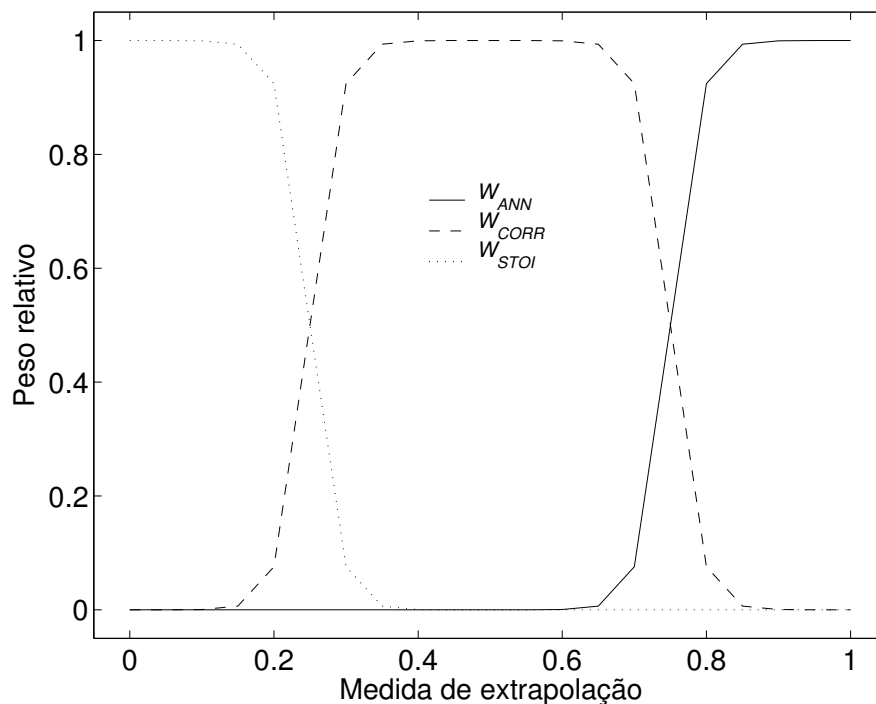


Figura 2.9: Peso relativo do modelo competitivo em função da medida de extrapolação.

É importante realçar que nenhum dos métodos discutidos faz a avaliação do desempenho real - exactidão - dos modelos em regiões diferentes do espaço das entradas e nenhum deles inclui essa informação na estratégia de ponderação.

2.5 Conclusões

Neste capítulo efectuou-se uma revisão das principais estruturas híbridas para modelização de processos (bio)químicos.

Os modelos híbridos, porque integram vários tipos de conhecimento, deverão ser classificados como modelos de caixa cinzenta. A estratégia de desenvolvimento é uma estratégia 'modular', a qual se rege pela identificação de subsistemas e do conhecimento disponível sobre cada subsistema. As estruturas resultantes podem ser classificadas como em série, paralelo, ou ambas. São sistemas potencialmente muito complexos. Existe ainda uma necessidade de desenvolver uma teoria adequada que suporte estes estudos,

por exemplo, sobre a identificação e a estabilidade. Só assim se poderá aumentar o potencial de aplicação prática desta técnica a processos industriais. Uma questão em aberto é a do método de ponderação em estruturas paralelas (ou competitivas). Este tema será abordado com profundidade neste trabalho de doutoramento.

Referências

- Anderson, J. S., McAvoy, T. J. e Hao, O. J. (2000), 'Use of hybrid models in wastewater systems', *Industrial & Engineering Chemistry Research*, **39**(6), 1694–1704.
- Babuska, R., Verbruggen, H. B. e van Can, H. J. L. (1999), 'Fuzzy modeling of enzymatic penicillin-g conversion', *Engineering Applications Of Artificial Intelligence*, **12**(1), 79–92.
- Bhat, N. e Mcavoy, T. J. (1990), 'Use of neural nets for dynamic modeling and control of chemical process systems', *Computers & Chemical Engineering*, **14**(4-5), 573–583.
- Braake, H. A. B. T., van Can, H. J. L. e Verbruggen, H. B. (1998), 'Semi-mechanistic modeling of chemical processes with neural networks', *Engineering Applications Of Artificial Intelligence*, **11**(4), 507–515.
- Chen, L., Bernard, O., Bastin, G. e Angelov, P. (2000), 'Hybrid modelling of biotechnological processes using neural networks', *Control Engineering Practice*, **8**(7), 821–827.
- Chinrungrueng, C. (1993), *Evaluation of heterogenous Architectures for Artificial Neural Networks*, Tese de Doutorado, University of California at Berkeley.
- Chinrungrueng, C. e Sequin, C. H. (1995), 'Optimal adaptive k-means algorithm with dynamic adjustment of learning rate', *Ieee Transactions On Neural Networks*, **6**(1), 157–169.
- Eikens, B. e Karim, M. N. (1999), 'Process identification with multiple neural network models', *International Journal Of Control*, **72**(7-8), 576–590.
- Feyo de Azevedo, S., Dahm, B. e Oliveira, F. R. (1997), 'Hybrid modelling of biochemical processes: A comparison with the conventional approach', *Computers & Chemical Engineering*, **21**, S751–S756.

- Feyo de Azevedo, S., Oliveira, R. e Sonnleitner, B. (2001), *Novel Multiphase Bioreactors*, Harwood Academic Publishers, UK, Capítulo 3: New Metodologies for Multiphase Bioreactors: Data Acquisition, Modelling and Control.
- Georgieva, P., Meireles, M. J. e de Azevedo, S. F. (2003), 'Knowledge-based hybrid modelling of a batch crystallisation when accounting for nucleation, growth and agglomeration phenomena', *Chemical Engineering Science*, **58**(16), 3699–3713.
- Horiuchi, J. e Hiraga, K. (1999), 'Industrial application of fuzzy control to large-scale recombinant vitamin b-2 production', *Journal Of Bioscience And Bioengineering*, **87**(3), 365–371.
- Johansen, T. A. e Foss, B. A. (1997), 'Operating regime based process modeling and identification', *Computers & Chemical Engineering*, **21**(2), 159–176.
- Karama, A., Bernard, O., Genovesi, A., Dochain, D., Benhammou, A. e Steyer, J. P. (2001a), 'Hybrid modelling of anaerobic wastewater treatment processes', *Water Science and Technology*, **43**(1), 43–50.
- Karama, A., Bernard, O., Gouze, J. L., Benhammou, A. e Dochain, D. (2001b), 'Hybrid neural modelling of an anaerobic digester with respect to biological constraints', *Water Science and Technology*, **43**(7), 1–8.
- Kohonen, T. (1995), *Self-Organizing Maps*, Springer, Heidelberg.
- Lauret, P., Boyer, H. e Gatina, J. (2000), 'Hybrid modelling of a sugar boiling process', *Control Engineering Practice*, **8**, 299–310.
- Lee, D. S., Vanrolleghem, P. A. e Park, J. M. (2005), 'Parallel hybrid modeling methods for a full-scale cokes wastewater treatment plant', *Journal Of Biotechnology*, **115**(3), 317–328.
- Lee, S. L., Jeon, C. O., Park, J. M. e Chang, K. S. (2002), 'Hybrid neural network modeling of a full-scale industrial wastewater treatment process', *Biotechnology and Bioengineering*, **78**(6), 670–682.
- Leonard, J. A., Kramer, M. A. e Ungar, L. H. (1992), 'A neural network architecture that computes its own reliability', *Computers & Chemical Engineering*, **16**(9), 819–835.
- Molga, E. J. (2003), 'Neural network approach to support modelling of chemical reactors: problems, resolutions, criteria of application', *Chemical Engineering And Processing*, **42**(8-9), 675–695.

- Montague, G. e Morris, J. (1994), 'Neural-network contributions in biotechnology', *Trends in Biotechnology*, **12**(8), 312–324.
- Oliveira, R. (1998), *Supervision, Control and Optimization of Biotechnological Processes Based on Hybrid Models*, Tese de Doutorado, Martin-Luther-Universitat Halle-Wittenberg.
- Oliveira, R., Peres, J. e Feyer de Azevedo, S. (2000), 'Efficient knowledge integration methods for improved bioreactor operation', em *4th Portuguese Conference on Automatic Control (Controlo'2000)*, Guimarães, Portugal, Páginas 214–218.
- Psichogios, D. C. e Ungar, L. H. (1992), 'A hybrid neural network-1st principles approach to process modeling', *AIChE Journal*, **38**(10), 1499–1511.
- Roubos, J. A., Krabben, P., Setness, M., Babuska, R., Heijnen, J. e Verbruggen, H. B. (1999), 'Hybrid model development for fed-batch bioprocesses combining physical equations with the metabolic network and black-box kinetics', em *6th Workshop on fuzzy systems*, Brunel University, Uxbridge, Páginas 231–239.
- Schubert, J., Simutis, R., Dors, M., Havlik, I. e Lubbert, A. (1994a), 'Bioprocess optimization and control application of hybrid modeling', *Journal of Biotechnology*, **35**(1), 51–68.
- Schubert, J., Simutis, R., Dors, M., Havlik, I. e Lubbert, A. (1994b), 'Hybrid modeling of yeast production processes combination of a-priori knowledge on different levels of sophistication', *Chemical Engineering & Technology*, **17**(1), 10–20.
- Senger, R. S. e Karim, M. N. (2003), 'Neural-network-based identification of tissue-type plasminogen activator protein production and glycosylation in CHO cell culture under shear environment', *Biotechnology Progress*, **19**(6), 1828–1836.
- Simutis, R., Havlik, I. e Lubbert, A. (1993), 'Fuzzy-aided neural network for real-time state estimation and process prediction in the alcohol formation step of production-scale beer brewing', *Journal of Biotechnology*, **27**(2), 203–215.
- Simutis, R., Havlik, I., Schneider, F., Dors, M. e Lübbert, A. (1995), 'Artificial neural networks of improved reliability for industrial process supervision', em *Preprints of the 6th Int. Conference on Computer Applications in Biotechnology*, Garmisch-Partenkirchen, Germany, Páginas 59–65.
- Simutis, R., Oliveira, R., Manikowski, M., de Azevedo, S. F. e Lubbert, A. (1997), 'How to increase the performance of models for process optimization and control', *Journal of Biotechnology*, **59**(1-2), 73–89.

- Thompson, M. L. e Kramer, M. A. (1994), 'Modeling chemical processes using prior knowledge and neural networks', *AIChE Journal*, **40**(8), 1328–1340.
- van Can, H. J. L., Braake, H. A. B. T., Hellinga, C., Luyben, K. C. A. M. e Heijnen, J. J. (1999), 'An efficient model development strategy for bioprocesses based on neural networks in macroscopic balances: Part ii', *Biotechnology and Bioengineering*, **62**(6), 666–680.
- van Can, H. J. L., teBraake, H. A. B., Dubbelman, S., Hellinga, C., Luyben, K. C. A. M. e Heijnen, J. J. (1998), 'Understanding and applying the extrapolation properties of serial gray-box models', *AIChE Journal*, **44**(5), 1071–1089.
- Weigend, A. S., Mangeas, M. e Srivastava, A. N. (1995), 'Nonlinear gated experts for time series: Discovering regimes and avoiding overfitting', *International Journal of Neural Systems*, **6**(4), 373–399.
- Zhao, H., Hao, O. J., McAvoy, T. J. e Chang, C. (1997), 'Modeling nutrient dynamics in sequencing batch reactor', *Journal of Environmental Engineering*, **123**, 311–319.

Capítulo 3

Descrição de Casos de Estudo

Conteúdo do Capítulo

Neste capítulo faz-se a descrição dos casos de estudo utilizados neste trabalho para testar os métodos propostos, a saber: produção de proteína recombinante em modo semicontínuo em culturas de levedura *Saccharomyces cerevisiae* da estirpe SEY2102-s2I, processo de produção de fermento de padeiro (levedura *Saccharomyces cerevisiae* da estirpe H1022 (ATCC 32167)), processo de remoção de fósforo de águas residuais por lamas activadas e processo de produção de Polihidroxialcanoatos. São apresentados modelos matemáticos que descrevem estes processos.

3.1 Caso de Estudo I: produção de proteína recombinante em modo semicontínuo

A produção de proteínas constitui uma das importantes aplicações da Engenharia Genética, existindo sistemas de expressão adequados à superprodução de proteínas recombinantes (codificadas por rDNA) desde as presentes em células bacterianas até às de mamífero (Sá-Correia et al., 2003).

A manipulação genética por mutação e selecção tem tido um papel importante no melhoramento dos níveis de expressão num conjunto muito grande de proteínas. No entanto, este melhoramento é obtido empiricamente, por tentativa e erro, tendo o investigador pouco controlo nas alterações genéticas adquiridas. A utilização da técnica

do DNA recombinante permite uma alteração direccionada de forma a obter alterações genéticas específicas, resultando na adição de novas propriedades e na ampliação das capacidades dos microrganismos industriais.

A expressão de proteínas recombinantes em células onde essa expressão não ocorre naturalmente é designada por produção de proteínas heterólogas. A utilização de bactérias e fungos como células hospedeiras na produção de proteínas heterólogas é prática corrente na indústria farmacêutica.

Contrariamente às proteínas sintetizadas intracelularmente, a produção de proteínas que são excretadas tornam mais simples a recolha e a purificação final das proteínas, evitando-se o rompimento das células. A taxa de secreção depende fortemente da actividade metabólica da célula hospedeira.

Walsh (2002) aponta algumas vantagens na produção de enzimas industriais pela técnica recombinante:

- são obtidos maiores níveis de expressão;
- o produto obtido apresenta duma maneira geral um grau de pureza relativo maior;
- é atractivo do ponto de vista económico;
- a expressão heteróloga facilita a comercialização de enzimas produzidas naturalmente por espécies patogénicas;
- permite a alteração das características das enzimas através de engenharia de proteínas.

Além destas vantagens, existem outras de natureza técnica e económica na super-produção de enzimas recombinantes, nomeadamente (Walsh, 2002):

- fermentações descontínuas de dimensão substancialmente menor com custos associados ao pré-processamento e pós-processamento menores;
- quantidades menores de resíduos e co-produtos que resultam de volumes de processamento menores;
- níveis de expressão elevados conduzem a taxas menores de contaminação das proteínas.

Park e Ramirez (1989) desenvolveram um modelo para descrever a dinâmica da secreção de proteínas heterólogas por células da levedura *Saccharomyces cerevisiae* da estirpe SEY2102-s2I, baseada nas principais interações macromoleculares entre os polipeptídeos e a maquinaria de secreção da célula hospedeira. Este modelo entra em consideração com a complexidade das reacções enzimáticas e com os mecanismos de transporte através da via de secreção.

Concretamente, após a síntese da proteína pelo ribossoma existente na face externa do retículo endoplasmático rugoso, a proteína atravessa a membrana do retículo endoplasmático (ER) onde é processada por enzimas e formatada correctamente na sua forma tridimensional. Após passar o ER, a proteína entra no complexo de Golgi onde é novamente modificada pela adição de cadeias de polissacáridos e empacotada dentro de vacúolos. O destino final da proteína é ser excretada para o meio extracelular por exocitose dos vacúolos que contém as proteínas (Lodish et al., 2000).

Park e Ramirez (1989) escolheram a levedura *S. cerevisiae* como hospedeiro pelas razões a seguir explicitadas:

- A atractividade desta levedura para a produção comercial de proteínas recombinantes advém das elevadas velocidades de crescimento, da facilidade de introdução de genes exógenos e pelo facto de não existir qualquer relação patogénica entre esta levedura e os humanos (Park e Ramirez, 1989). Com efeito, na indústria farmacêutica a purificação e a remoção de toxinas é importante o que torna a *S. cerevisiae* num microrganismo atractivo para produzir enzimas terapêuticas e neuropeptídeos (Park e Ramirez, 1988).
- A capacidade das células de levedura para excretar proteínas com modificações pós-tradução específicas tais como a glicosilação e a fosforilação (Park e Ramirez, 1988) torna o fermento de padeiro num organismo com potencial para a superprodução de proteínas humanas que requeiram este tipo de modificações. De facto, a maquinaria de secreção das células de levedura é bastante análoga à maquinaria de secreção das células dos mamíferos.

De seguida descreve-se o modelo dinâmico adoptado neste trabalho, como caso de estudo, proposto por Park e Ramirez (1988).

A dinâmica do processo é descrito por um conjunto de balanços materiais em reactor

semicontínuo, onde a taxa de alimentação, F , é a única acção de controlo:

$$\frac{dX}{dt} = \mu(S)X - DX \quad (3.1)$$

$$\frac{dS}{dt} = -\frac{1}{Y}\mu(S)X + D(S_o - S) \quad (3.2)$$

$$\frac{dP_t}{dt} = f_p(S)X - DP_t \quad (3.3)$$

$$\frac{dP_m}{dt} = \Phi(S)(P_t - P_m) - DP_m \quad (3.4)$$

em que X é a concentração de biomassa, S é a concentração de glucose, P_t é a concentração total de proteína por unidade de volume do reactor, e portanto inclui proteína já excretada e aquela proteína que ainda não foi excretada, P_m é a concentração de proteína excretada para o meio de cultura por unidade de volume do reactor, D é a taxa de diluição ($D = F/V$ sendo F a taxa de alimentação e V o volume do meio dentro do biorreactor), S_o a concentração de substrato na corrente de entrada, μ é a taxa específica de crescimento, f_p é a taxa de expressão de proteína e Φ é a taxa de excreção de proteína.

Uma vez que o processo é operado em modo semicontínuo e assumindo que a densidade do meio não é constante, a equação de balanço material global

$$\frac{dV}{dt} = F - DV \quad (3.5)$$

tem de ser integrada juntamente com as Equações (3.1)-(3.4).

Assume-se que o coeficiente de rendimento massa celular/glucose, Y , é constante igual a 1/7.3 e a concentração de glucose na alimentação também é constante. Sob condições isotérmicas de operação, as taxas cinéticas μ , f_p e Φ dependem apenas da concentração de glucose na cultura e são descritas pelas seguintes equações:

$$\mu(S) = \frac{21.87S}{(S + 0.4)(S + 62.5)} \quad (3.6)$$

$$f_p(S) = \frac{Se^{-5S}}{(S + 0.1)} \quad (3.7)$$

$$\Phi(S) = \frac{4.75\mu(S)}{0.12 + \mu(S)} \quad (3.8)$$

em que S representa a concentração de glucose expressa em g/L, μ representa a taxa específica de crescimento expressa em 1/h, f_p representa a taxa de expressão de proteína expressa em u.a./(g biomassa.h) e Φ representa a taxa de excreção de proteína expressa

em 1/h.

Este modelo foi validado experimentalmente com a estirpe de levedura SEY2102-s2I em Park e Ramirez (1988).

3.2 Caso de Estudo II: processo de produção de fermento de padeiro

O fermento de padeiro utilizado nas indústrias de panificação para fazer levedar a massa, é composto, essencialmente por células vivas da levedura *Saccharomyces cerevisiae*. As leveduras são microrganismos heterotróficos, isto é, são organismos que podem obter energia e fonte de carbono a partir de compostos orgânicos. Daí que o crescimento destes microrganismos dependa de uma variedade de compostos orgânicos e alguns nutrientes minerais (Reed e Peppler, 1973). As leveduras são microrganismos anaeróbios facultativos, isto é, tanto crescem na ausência de oxigénio, onde ocorre a fermentação do substrato, como na sua presença, onde neste caso ocorre a respiração e o metabolismo oxidativo do substrato.

Nas secções seguintes apresenta-se o modelo matemático de simulação adoptado para descrever a produção de fermento de padeiro e descreve-se sucintamente a origem dos dados experimentais utilizados.

3.2.1 Modelo de simulação da produção de fermento de padeiro

Adoptou-se o modelo cinético proposto por Sonnleitner e Kappeli (1986) para o crescimento da levedura *Saccharomyces cerevisiae*, já utilizado em estudos anteriores por Oliveira (1997) e Ferreira (1995). Este modelo, é baseado no princípio da capacidade respiratória limitada. O modelo considera três vias metabólicas (ver Figura 3.1) para a utilização da fonte de carbono que podem ser juntas em três reacções macroscópicas com estequiometria bem definida:



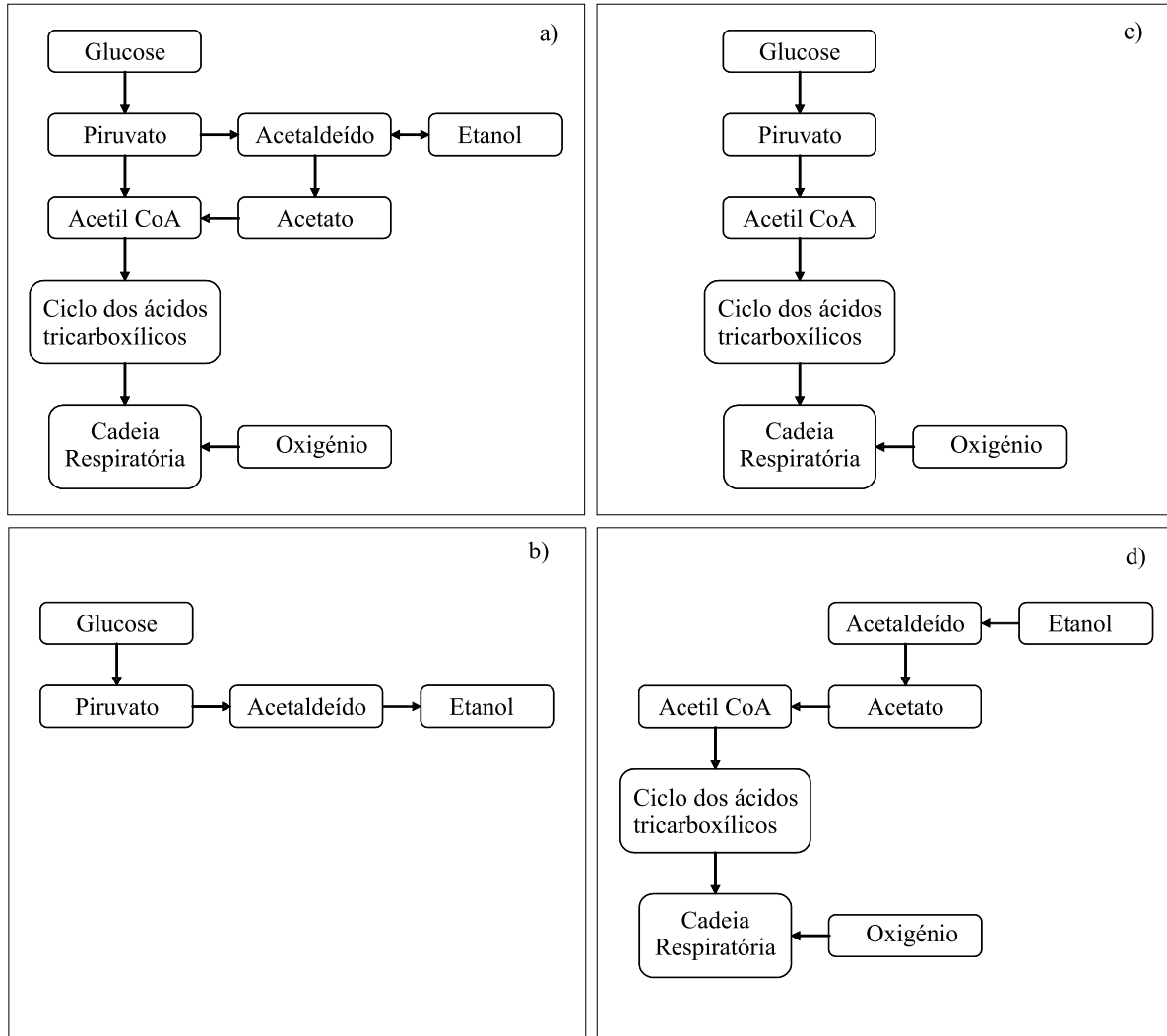


Figura 3.1: Representação esquemática simplificada das 3 vias metabólicas para o crescimento da *S. cerevisiae* em glucose e etanol (a) metabolismo global para o crescimento em glucose (b) fermentação da glucose (c) oxidação da glucose (d) oxidação do etanol (adaptado de Oliveira, 1997)

sendo X a concentração de biomassa, S a concentração de glucose, E a concentração de etanol, N a concentração de amónia, O_2 a concentração de oxigénio, CO_2 a concentração de dióxido de carbono e μ_s^o , μ_s^r e μ_e^o são as três taxas específicas de crescimento associadas a cada uma das reacções.

A taxa específica de crescimento total, μ , é igual à soma das três taxas específicas de crescimento referidas, ou seja,

$$\mu = \mu_s^o + \mu_s^r + \mu_e^o \quad (3.12)$$

As vias metabólicas do crescimento oxidativo em glucose são governadas pela capaci-

dade respiratória das células. Apenas ocorrerá o metabolismo oxidativo se existir glucose em concentrações pequenas e houver oxigénio suficiente disponível no meio. Neste caso, a glucose é o substrato preferencial em vez do etanol. No entanto, se o fluxo de glucose exceder a capacidade respiratória máxima, uma parte é catabolizado oxidativamente e o restante seguirá o catabolismo fermentativo havendo produção de etanol. Esta situação corresponde ao estado oxido-redutivo. Na presença de baixas concentrações de glucose o etanol formado pela via redutiva pode ser consumido por via oxidativa.

A taxa específica (total) de consumo de glucose, q_s , pode ser expressa pelo modelo cinético de MONOD:

$$q_s = q_{smax} \frac{S}{K_s + S} \quad (3.13)$$

em que q_{smax} é o valor máximo para a taxa específica de consumo de glucose e K_s é a constante de saturação.

A utilização de etanol é influenciada pela prioridade do consumo de glucose, a qual funciona como inibidor. A taxa específica de crescimento em etanol pode ser descrita pelo valor mínimo entre as duas taxas μ_{e1}^o e μ_{e2}^o definidas pelas expressões seguintes:

$$\mu_{e1}^o = \mu_{emax} \frac{E}{K_e + E} \frac{K_i}{K_i + S} \quad (3.14)$$

$$\mu_{e2}^o = Y_{xo}^{oe} \left(q_{omax} - q_s \frac{Y_{xs}^o}{Y_{xo}^o} \right) \quad (3.15)$$

em que μ_{emax} é o valor máximo para a taxa específica de crescimento em etanol, K_e é a constante de saturação e K_i é uma constante de inibição, q_{omax} é a taxa específica máxima de consumo de oxigénio, Y_{xo}^{oe} é o coeficiente de rendimento de biomassa em etanol na fase oxidativa em etanol, Y_{xs}^o e Y_{xo}^o representam os coeficientes de rendimento de biomassa em glucose e da biomassa em oxigénio, respectivamente. A equação (3.14) impõe uma restrição tal que a oxidação do etanol só ocorre se existir capacidade respiratória disponível.

A quantidade de oxigénio necessária para oxidar a glucose é $q_s Y_{xs}^o / Y_{xo}^o$. Assim sendo, duas situações podem ocorrer:

1. existência de oxigénio em excesso, isto é, quantidade de oxigénio existente no meio superior à necessária pra oxidar q_s , ocorrendo as vias oxidativas em glucose e etanol. Neste caso o crescimento fermentativo é nulo;
2. existência de oxigénio inferior ao necessário para oxidar q_s , o que implica a satura-

ção do meio em glucose, ocorrendo a via oxido-redutiva em glucose. Neste caso, o crescimento oxidativo em etanol é nulo.

Portanto, em cada instante, ocorrem somente duas das três vias metabólicas. A Tabela 3.1 concretiza matematicamente o exposto assumindo que o processo ocorre em aerobiose e que o oxigénio está sempre em grande excesso.

Tabela 3.1: Taxas específicas de crescimento do fermento de padeiro

via oxidativa	via oxido-redutiva
$q_s \leq q_{omax} Y_{xo}^o / Y_{xs}^o$	$q_s > q_{omax} Y_{xo}^o / Y_{xs}^o$
$\mu_s^o = Y_{xs}^o q_s$	$\mu_s^o = Y_{xo}^o q_{omax}$
$\mu_s^r = 0$	$\mu_s^r = Y_{xs}^r (q_s - q_{omax} Y_{xo}^o / Y_{xs}^o)$
$\mu_e^o = \min(\mu_{e1}^o, \mu_{e2}^o)$	$\mu_e^o = 0$

Utilizaram-se para valores dos parâmetros cinéticos os valores apresentados na Tabela 3.2.

Tabela 3.2: Parâmetros cinéticos (tirados de Sonnleitner e Kappeli, 1986)

$q_{smax} = 3.5 \text{ g glucose / (g biomassa.h)}$	$K_s = 0.2 \text{ g/L}$
$\mu_{emax} = 0.17 \text{ h}^{-1}$	$K_e = 0.1 \text{ g/L}$
$q_{omax} = 0.256 \text{ g O}_2\text{/ (g biomassa.h)}$	$K_i = 0.1 \text{ g/L}$

O modelo dinâmico para o crescimento do fermento de padeiro num biorreactor semicontínuo é obtido por balanços materiais aos componentes intervenientes no esquema reaccional (3.9) - (3.11) considerando que o reactor é perfeitamente agitado, que os coeficientes de rendimento são constantes e que a dinâmica da fase gasosa pode ser desprezada. Os balanços materiais são expressos pelas equações seguintes:

$$\frac{dX}{dt} = (\mu_s^o + \mu_s^r + \mu_e^o - D) X \quad (3.16)$$

$$\frac{dS}{dt} = - \left(\frac{\mu_s^o}{Y_{xs}^o} + \frac{\mu_s^r}{Y_{xs}^r} \right) X - D(S - S_o) \quad (3.17)$$

$$\frac{dE}{dt} = \left(\frac{\mu_s^r}{Y_{xe}^r} - \frac{\mu_e^o}{Y_{xe}^{oe}} \right) X - DE \quad (3.18)$$

$$\frac{dO}{dt} = - \left(\frac{\mu_s^o}{Y_{xo}^o} + \frac{\mu_e^o}{Y_{xo}^{oe}} \right) X - DO + OTR \quad (3.19)$$

$$\frac{dC}{dt} = \left(\frac{\mu_s^o}{Y_{xc}^o} + \frac{\mu_s^r}{Y_{xc}^r} + \frac{\mu_e^o}{Y_{xc}^{oe}} \right) X - DC - CTR \quad (3.20)$$

e a equação adicional de balanço global

$$\frac{dV}{dt} = DV = F, \quad (3.21)$$

em que D é a taxa de diluição, Y_i^j são coeficientes de rendimento, S_o é a concentração de glucose na alimentação, CTR é a taxa de transferência de dióxido de carbono e OTR é a taxa de transferência de oxigénio.

Na Tabela 3.3 apresentam-se os valores referentes aos coeficientes de rendimento.

Tabela 3.3: Rendimentos estequiométricos (tirados de Pomerleau e Perrier, 1990)

$Y_{xs}^o (= 1/k_1) = 0.49$	g biomassa / (g glucose)
$Y_{xs}^r (= 1/k_2) = 0.05$	g biomassa / (g glucose)
$Y_{xe}^r (= 1/k_3) = 0.1$	g biomassa / (g etanol)
$Y_{xe}^{oe} (= 1/k_4) = 0.72$	g biomassa / (g etanol)
$Y_{xo}^o (= 1/k_5) = 1.2$	g biomassa / (g oxigénio)
$Y_{xo}^{oe} (= 1/k_6) = 0.64$	g biomassa / (g oxigénio)
$Y_{xc}^o (= 1/k_7) = 0.81$	g biomassa / (g dióxidoC)
$Y_{xc}^r (= 1/k_8) = 0.11$	g biomassa / (g dióxidoC)
$Y_{xc}^{oe} (= 1/k_9) = 1.11$	g biomassa / (g dióxidoC)

3.2.2 Dados experimentais

As experiências foram efectuadas (Oliveira, 1997) em laboratório com a levedura *Saccharomyces cerevisiae* e a estirpe utilizada foi a H1022 (ATCC 32167) num fermentador com capacidade máxima de 5 L. O meio de cultura semi-sintético utilizado nas fermentações está indicado na Tabela 3.4. O meio é esterilizado em autoclave a 121°C durante 20-30 minutos.

Tabela 3.4: Composição do meio de cultura para produção de fermento de padeiro

Composto	Concentração do Inóculo (g/L)	Concentração do Meio Inicial (g/L)
Açúcares	5 ou 30	1 ou 5
KH_2PO_4	5	5
$(NH_4)_2SO_4$	2	2
$MgSO_4 \cdot 7H_2O$	0.4	0.4
Extracto de levedura	1	1

As condições experimentais no decorrer de cada experiência são mantidas constantes dentro dos valores de referência indicados na Tabela 3.5.

Tabela 3.5: Condições experimentais

Variável	Valor de referência
Temperatura	30° C
pH	4.0
Velocidade de agitação	500 rpm
Arejamento	3.5 slpm

Na Tabela 3.6 apresentam-se as condições de operação de cada experiência onde $X(0)$, $S(0)$ e $E(0)$ são os valores iniciais da biomassa, da glucose e do etanol, respectivamente. O volume inicial $V(0)$ é constante. F é a taxa de alimentação e S_o é a concentração de glucose na alimentação.

Tabela 3.6: Condições da fermentação

Partida	$X(0)(g/L)$	$S(0)(g/L)$	$E(0)(g/L)$	$V(0)(L)$	$F(L/h)$	$S_o(g/L)$
B1	1.20	1.46	2.27	2.5	0.12	50
B2	1.54	0.29	2.90	2.5	0.15	100
B3	0.38	1.53	1.95	2.5	0.15	50
B4	1.46	0.00	1.84	2.5	0.15	25
B5	0.23	3.13	0.72	2.5	0.05	10
B6	0.25	2.98	0.64	2.5	0.10	25
B7	1.40	25.44	3.80	2.5	0.15	5

3.3 Caso de Estudo III: processo de remoção de fósforo de águas residuais por lamas activadas

O processo de remoção de fósforo de águas residuais provenientes da agricultura, de efluentes industriais e domésticos é um processo determinante para controlar um dos mais sérios problemas ambientais actuais: a eutrofização. Com efeito, além do fósforo, o carbono e o azoto em excesso (mais o azoto do que o carbono) também são responsáveis pelo crescimento desmesurado de algas e plantas em meios aquáticos. No entanto, se o elemento limitante for o azoto, as algas cianofícias, fixadoras de azoto atmosférico, fornecem ao meio aquático este composto. Sendo assim, dever-se-á controlar o fós-

foro reduzindo o teor de fosfatos nas águas residuais de modo a tornar este composto limitante.

Os processos habitualmente utilizados para a remoção de fósforo de águas residuais são a precipitação química e a remoção biológica. A remoção biológica tem a vantagem de, sem custos adicionais, se obterem efluentes com concentrações muito baixas de fósforo. Além disso, pode conjugar-se a remoção biológica de fósforo com a remoção biológica de nitratos. Actualmente, a remoção biológica de fósforo é efectuada exclusivamente por lamas activadas.

Os microrganismos responsáveis pela remoção de concentrações elevadas de fósforo de efluentes são as Bactérias Acumuladoras de Fósforo (*Phosphorous Accumulating Organisms* - PAOs). Nas estações de tratamento biológico de efluentes contaminados com fósforo, a biomassa recircula continuamente entre ambientes anaeróbios e aeróbios o que estimula a síntese de reservas intracelulares, nomeadamente, de Polihidroxialcanoatos (PHAs), polifosfatos e glicogénio.

Em condições anaeróbias as PAOs consomem substratos orgânicos, como por exemplo acetato, que são armazenados intracelularmente na forma de PHAs. A energia (ATP) e os redutores equivalentes (NADH) necessários neste processo, são gerados pela degradação das reservas internas de polifosfato e glicogénio, respectivamente. Na fase aeróbia, onde ocorre a remoção do fósforo, os PHAs são degradados para crescimento celular, para síntese de polifosfato e produção de glicogénio. O polifosfato formado nesta fase, resulta da polimerização dos fosfatos existentes no meio extracelular, que assim são removidos do meio.

O modelo matemático de simulação adoptado neste trabalho é baseado no modelo das lamas activadas n°2d (ASM2d, Henze et al., 1999). Este modelo é um modelo estruturado complexo que entra em consideração com a existência e a interacção de 3 grupos de microrganismos, nomeadamente, bactérias heterotróficas, bactérias de acumulação de fósforo e bactérias autotróficas. O modelo foi simplificado para um único grupo de microrganismos, nomeadamente as bactérias de acumulação de fósforo.

Os componentes que entram no modelo simplificado são mencionados na Tabela 3.7.

A fracção, f_i , do componente i é definido relativamente à biomassa activa, X_{PAO} , pela seguinte expressão:

$$f_i = \frac{X_i}{X_{PAO}} \quad (3.22)$$

Tabela 3.7: Componentes considerados no modelo simplificado

Componente	Descrição
S_{O_2}	concentração de oxigénio dissolvido
S_F	concentração de substrato fermentável
S_A	concentração de acetato
S_{NH_4}	concentração de amónia
S_{PO_4}	concentração de fosfato
S_I	fracção inerte
S_{ALK}	alcalinidade em bicarbonato
X_S	concentração de substrato lentamente biodegradável
X_{PAO}	concentração de biomassa activa, isto é, de PAOs
X_{PP}	concentração de polifosfato armazenado (intracelular)
X_{PHA}	concentração PHA armazenado (intracelular)
X_{GLY}	concentração de glicogénio armazenado (intracelular)

Seguidamente apresentam-se as equações cinéticas associadas às reacções metabólicas consideradas:

1. hidrólise em aerobiose:

$$r_1 = k_h \frac{f_S}{K_X + f_S} \frac{S_{O_2}}{K_{O_2}^L + S_{O_2}} X_{PAO} \quad (3.23)$$

2. hidrólise em anaerobiose:

$$r_3 = \eta_{fe}^L k_h \frac{f_S}{K_X + f_S} \frac{K_{O_2}^L}{K_{O_2}^L + S_{O_2}} X_{PAO} \quad (3.24)$$

3. acumulação de acetato na forma de PHA em anaerobiose:

$$r_{20} = q_{S,AN}^{max} \frac{S_A}{K_A + S_A} \frac{f_{PHA}^{max} - f_{PHA}}{K_{fPHA} + f_{PHA}^{max} - f_{PHA}} \frac{X_{GLY}}{K_{GLY} + X_{GLY}} \frac{X_{PP}}{K_{PP} + X_{PP}} X_{PAO} \quad (3.25)$$

4. manutenção em anaerobiose:

$$r_{21} = m_{AN} \frac{K_{O_2}}{K_{O_2} + S_{O_2}} \frac{X_{PP}}{K_{PP} + X_{PP}} X_{PAO} \quad (3.26)$$

5. consumo de PHA em aerobiose:

$$r_{22} = k_{PHA} \frac{f_{PHA}}{K_{fPHA} + f_{PHA}} \frac{S_{O_2}}{K_{O_2} + S_{O_2}} \frac{S_{NH_4}}{K_{NH_4} + S_{NH_4}} \frac{S_{PO_4}}{K_P + S_{PO_4}} \frac{S_{ALK}}{K_{ALK} + S_{ALK}} X_{PAO} \quad (3.27)$$

6. acumulação de polifosfatos em aerobiose a partir de PHA:

$$r_{23} = k_{PP} \frac{X_{PAO}}{X_{PP}} \frac{S_{PO_4}}{K_{PO_4} + S_{PO_4}} \frac{S_{O_2}}{g_{PP} K_{O_2} + S_{O_2}} \frac{f_{PP}^{max} - f_{PP}}{K_{fPP} + f_{PP}^{max} - f_{PP}} \frac{X_{PHA}}{K_{PHA} + X_{PHA}} X_{PAO} \quad (3.28)$$

7. acumulação de glicogénio em aerobiose a partir de PHA:

$$r_{24} = k_{GLY} \frac{X_{PHA}}{X_{GLY}} \frac{S_{O_2}}{K_{O_2} + S_{O_2}} \frac{f_{GLY}^{max} - f_{GLY}}{K_{fGLY} + f_{GLY}^{max} - f_{GLY}} \frac{X_{PHA}}{K_{PHA} + X_{PHA}} X_{PAO} \quad (3.29)$$

8. manutenção em aerobiose:

$$r_{25} = m_{O_2} \frac{S_{O_2}}{K_{O_2} + S_{O_2}} X_{PAO} \quad (3.30)$$

Os valores dos parâmetros estequiométricos e dos parâmetros cinéticos estão na Tabela 3.8.

Neste trabalho, estuda-se a remoção de fósforo em reactor descontínuo sequencial (*Sequencing Batch Reactor* - SBR). Cada ciclo de operação do SBR consiste em duas fases de reacção. A primeira fase que é a anaerobiose, sendo imediatamente seguida pela aerobiose. A transição entre as fases anaerobiose e aerobiose é efectuada ligando ou desligando o arejamento.

As equações de balanço material aos componentes considerados no modelo simplificado são as seguintes:

$$\frac{dS_{O_2}}{dt} = \left(\frac{1}{Y_{PHA}^O} - 1 \right) r_{22} - \frac{1}{Y_{PP}^O} r_{23} + \left(1 - \frac{1}{Y_{GLY}^O} \right) r_{24} - r_{25} \quad (3.31)$$

$$\frac{dS_F}{dt} = (1 - f_{SI}) (r_1 + r_3) \quad (3.32)$$

Tabela 3.8: Parâmetros do processo

Parâmetros estequiométricos			
$f_{SI} = 0$	(g-COD/g-COD)	$Y_{PHA}^O = 1.39$	(g-COD/g-COD)
$Y_{PO4} = 0.35$	(g-P/g-COD)	$Y_{GLY}^O = 1.11$	(g-COD/g-COD)
$Y_{PHA} = 1.50$	(g-COD/g-COD)	$Y_{PP}^O = 4.42$	(g-P/g-COD)
Parâmetros cinéticos			
$k_h = 3.0$	(g-COD/g-COD.d)	$K_{O2}^L = 0.20$	(g- O_2 /m ³)
$\eta_{fe}^L = 0.2$		$K_X = 0.1$	(g-COD/g-COD)
$q_{S,AN}^{max,AP} = 8.0$	(g-COD/g-COD.d)	$k_{PHA} = 5.51$	(g-COD/g-COD.d)
$m_{AN} = 0.05$	(g-P/g-COD.d)	$k_{PP} = 0.10$	(g-P/g-COD.d)
$k_{GLY} = 0.93$	(g-COD/g-COD.d)	$g_{PP} = 0.22$	
$m_{O2} = 0.06$	(g- O_2 /g-COD.d)	$K_P = 1.00$	(g-P/m ³)
$K_A = 4.00$	(g-COD/m ³)	$K_{fPHA} = 0.20$	(g-COD/g-COD)
$K_{O2} = 0.20$	(g- O_2 /m ³)	$K_{PO4} = 0.02$	(g-P/m ³)
$K_{NH4} = 0.05$	(g-N/m ³)	$f_{PP}^{max} = 0.35$	(g-P/g-COD)
$f_{PHA}^{max} = 0.05$	(g-COD/g-COD)	$f_{GLY}^{max} = 0.50$	(g-COD/g-COD)
$K_{PP} = 0.01$	(g-P/m ³)	$K_{PHA} = 0.01$	(g-COD/m ³)
$K_{GLY} = 0.01$	(g-COD/m ³)	$K_{fGLY} = 0.01$	(g-COD/g-COD)
$K_{fPP} = 0.01$	(g-P/g-COD)	$K_{ALK} = 0.01$	(mol- HCO_3^- /m ³)

$$\frac{dS_A}{dt} = -r_{20} \quad (3.33)$$

$$\frac{dS_{PO4}}{dt} = Y_{PO4} r_{20} + r_{21} - 0.0144 r_{22} - 0.9955 r_{23} + 0.0180 r_{24} + 0.02 r_{25} \quad (3.34)$$

$$\frac{dX_S}{dt} = -r_1 - r_3 \quad (3.35)$$

$$\frac{dX_{PAO}}{dt} = \frac{1}{Y_{PHA}^O} r_{22} - \frac{1}{Y_{PP}^O} r_{23} - \frac{1}{Y_{GLY}^O} r_{24} - r_{25} \quad (3.36)$$

$$\frac{dX_{PP}}{dt} = -Y_{PO4} r_{20} - r_{21} + r_{23} \quad (3.37)$$

$$\frac{dX_{PHA}}{dt} = Y_{PHA} r_{20} - r_{22} \quad (3.38)$$

$$\frac{dX_{GLY}}{dt} = (1 - Y_{PHA}) r_{20} + r_{24} \quad (3.39)$$

Assume-se que S_{NH4} , S_{ALK} , S_I são constantes e portanto as equações de balanço material correspondentes são iguais a zero. Assume-se também que a concentração de amónia é suficientemente elevada de forma a nunca ser limitante.

3.4 Caso de Estudo IV: processo de produção de Polihidroxialcanoatos

O Poli- β -hidroxibutirato (PHB) é um polímero biodegradável com propriedades semelhantes ao polímero sintético polipropileno. O custo de produção de PHB é ainda muito superior ao do polipropileno. Este factor é o principal obstáculo à substituição do polipropileno pelo PHB. Segundo Serafim et al. (2004) os custos de produção de PHB podem ser substancialmente reduzidos com culturas mistas e substratos mais baratos.

Serafim et al. (2004) demonstraram que culturas mistas podem acumular até 78% (W/W) em peso de PHB intracelularmente num reactor descontínuo sequencial. O substrato utilizado foi o ácido acético. A acumulação intracelular de PHB em culturas mistas é controlada pela estratégia de alimentação da fonte de carbono. A acumulação de PHB ocorre quando as populações mistas são sujeitas a ciclos de 'fartura' e 'fome' (Beccari et al., 1998; Beun et al., 2002). Isto é, um período curto com grande excesso de substrato é alternado com um período longo de carência ou mesmo de ausência de substrato (ver o ciclo de 'fartura' e 'fome' representado na Figura 3.2). Durante a fase de 'fartura' o substrato consumido está directamente relacionado com a acumulação de PHB e em menor escala com o crescimento da biomassa (dependendo da alimentação da fonte de azoto). Quando o substrato é totalmente consumido (fase de 'fome'), o polímero acumulado é utilizado como fonte de energia e de carbono para manutenção e crescimento. A carência de carbono durante um período longo de tempo provoca

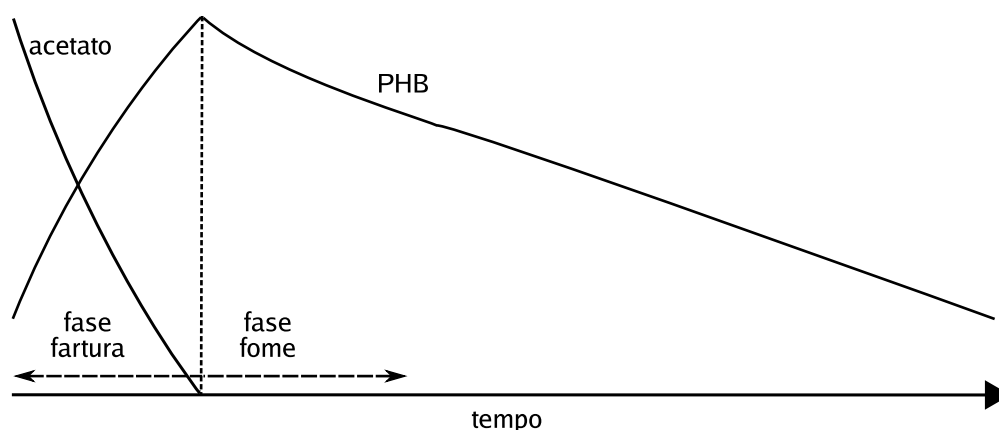


Figura 3.2: Ciclo de 'fartura' e 'fome'

alterações na composição macromolecular das células obrigando-as a uma adaptação fisiológica quando expostas a uma concentração elevada de substrato (Daigger e Grady, 1982).

Na literatura encontram-se descritos vários modelos matemáticos que tentam descrever os mecanismos de acumulação e degradação de PHB em culturas mistas. Estes estudos não tiveram como objectivo a optimização do processo, mas sim o estudo dos mecanismos subjacentes ao processo de acumulação de reservas internas. O modelo ASM3 tornou-se a referência para analisar o processo de lamas activadas. O modelo ASM3 entra em consideração com dois grupos de organismos (os organismos heterotróficos e os organismos autotróficos) e tenta descrever os processos de crescimento celular, de nitrificação e de desnitrificação, assim como, a acumulação de substratos orgânicos, tanto a acumulação aeróbia de COD como a acumulação anóxia de COD (Gujer et al., 1999).

O modelo ASM3 pode ser simplificado eliminando os organismos autotróficos obtendo-se, assim, um modelo que descreve apenas a acumulação aeróbia heterotrófica do acetato na forma de PHB. Este modelo simplificado tem uma desvantagem considerável pelo facto de não contemplar com a ocorrência em simultâneo do crescimento celular e da acumulação de PHB em COD. Segundo este modelo, o COD tem de ser inicialmente armazenado na forma de PHB e só depois pode ser metabolizado para crescimento.

Krishna e Van Loosdrecht (1999a,b) e Carucci et al. (2001) propuseram um modelo simplificado, baseado no modelo ASM3, descrevendo a conversão heterotrófica mas contemplando o crescimento e a acumulação de PHB em simultâneo. Este modelo simplificado implicou alterações significativas na estequiometria e nos parâmetros cinéticos do processo mas desta forma pode-se aumentar a exactidão do modelo.

Na literatura encontra-se ainda o estudo de outros modelos metabólicos, mais complexos, que tentam descrever a acumulação de PHB pelo processo das lamas activadas.

Beun et al. (2000) propõem um modelo baseado em 7 reacções metabólicas. Este modelo foi adaptado a partir dum modelo de *Paracoccus pantotrophus* (van Aalst-van Leeuwen et al., 1997) e melhorado posteriormente por Beun et al. (2000, 2002). O cálculo dos coeficientes de rendimento e manutenção teóricos é desenvolvido a partir de balanços materiais e energéticos das reacções metabólicas.

van Loosdrecht e Heijnen (2002) definiram um modelo estruturado para a biomassa onde se descreve a formação e a degradação de determinada enzima com o objectivo de descrever a modulação de substrato que conduz à formação de biomassa e à acumulação de polímero. O modelo demonstrou simular o comportamento típico dum sistema dinâmico contínuo mas não foi validado experimentalmente.

Third et al. (2003) apresentaram um modelo focado no efeito do oxigénio dissolvido

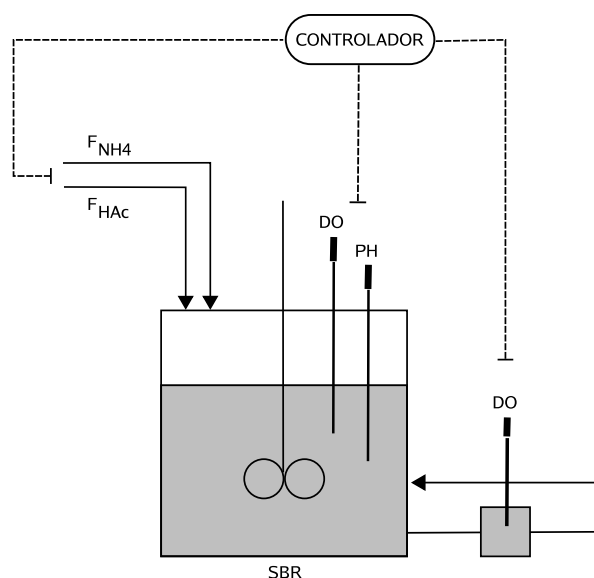


Figura 3.3: Esquema da operação do processo de produção de PHB por culturas mistas.

na conversão de COD em PHB. Este parece ser um factor importante na remoção de nitrogénio em estações de tratamento de águas residuais quando a razão entre a fonte de carbono e a amónia no efluente é baixa. O modelo considera 6 reacções metabólicas: consumo de acetato, respiração, crescimento celular, ciclo dos ácidos tricarboxílicos (TCA), formação e degradação de PHB. Foi efectuado um estudo de simulação que mostra que taxas altas de fornecimento de oxigénio favorecem o crescimento à custa duma redução das taxas de formação de PHB. Este modelo também não foi validado experimentalmente.

Em todos estes modelos se observa que a exactidão e capacidade predictiva não são suficientes para estudos quantitativos baseados em modelos (ver relatórios técnicos dos modelos ASM em Henze et al., 2000). Parece portanto haver uma janela de oportunidade para a técnica híbrida para a modelização e optimização destes processos.

Na Figura 3.3 apresenta-se um esquema do processo de produção de PHB a partir de ácido acético por culturas mistas num reactor descontínuo sequencial.

A actividade experimental vem descrita detalhadamente em Serafim et al. (2004). De uma forma resumida, as experiências foram desenhadas para diferentes razões de carbono/nitrogénio na alimentação. A alimentação de carbono (ácido acético) e de fonte de nitrogénio (amónia) foi controlada pela concentração de oxigénio dissolvido. Estes compostos são adicionados por pulsos. O consumo total de fonte de carbono é detectado por um aumento brusco de concentração do oxigénio dissolvido. Um algoritmo de estimativa recursivo detecta os picos de oxigénio e comanda a adição de novos pulsos

de substratos.

A temperatura do reactor foi mantida a 22° C, a velocidade de agitação a 250 rpm e o pH não foi controlado.

O meio de sais minerais standard é composto por (por litro de água destilada) 4.0 g CH₃COO Na.3H₂O (30 C-mmol), 600 mg MgSO₄.7H₂O, 160 mg NH₄Cl (1.4 N-mmol), 100 mg EDTA, 70 mg CaCl₂.2H₂O e 2 ml de traçador. A solução do traçador é constituída por (por litro de água destilada) 1500 mg FeCl₃.6H₂O, 150 mg H₃BO₃, 150 mg CoCl₂.6H₂O, 120 mg MnCl₂.4H₂O, 120 mg ZnSO₄.7H₂O, 60 mg Na₂MoO₄.2H₂O, 30 mg CuSO₄.5H₂O e 30 mg of KI. Para inibir a nitrificação foi adicionado Thiourea (10 mg/L). O pH da solução de sais minerais foi ajustado para 7.2 e posteriormente a solução foi esterilizada. Após a esterilização foi adicionado ao meio de cultura uma solução de fósforo composta por 92 mg K₂HPO₄ e 45 mg KH₂PO₄ por litro de água destilada.

Foram usadas diferentes concentrações de acetato e amónia nas experiências efectuadas (Serafim et al., 2004). Concretamente, para acetato: 15 C-mmol/L, 30 C-mmol/L, 60 C-mmol/L, 90 C-mmol/L, e 180 C-mmol/L e para a amónia: 0 N-mmol/L, 0.7 N-mmol/L, 1.4 N-mmol/L, e 2.8 N-mmol/L.

Foram efectuadas medidas das variáveis que definem o estado do processo, nomeadamente as concentrações da biomassa activa, X , do acetato, HAc , de amónia, NH_4 e do produto poli- β -hidroxibutirato, PHB .

3.5 Conclusões

Neste capítulo descreveram-se os processos biológicos que servirão como casos de estudo em capítulos futuros. Apresentaram-se para todos os processos biológicos um modelo matemático detalhado do processo exceptuando o de produção de PHAs (dado que não existe actualmente nenhum modelo para os dados experimentais considerados). Estes modelos serão utilizados em capítulos seguintes para gerar dados e assim permitir a validação de detalhes específicos de modelização híbrida de bioprocessos.

De realçar também, o facto de estes processos e respectivos modelos exibirem níveis de complexidade distintos. No caso I, o processo é descrito por um modelo dinâmico altamente não linear mas as cinéticas dependem de uma única variável de estado. No caso II, o processo é descrito por um modelo não-estruturado, mas onde estão envolvidas três reacções, podendo ocorrer trocas entre elas. Nos casos III e IV, os sistemas biológicos

acumulam reservas intracelulares, pelo que a modelização dinâmica do processo terá que obrigatoriamente considerar a própria dinâmica dos componentes da fase intracelular. Estes diferentes níveis de complexidade, poderão portanto proporcionar um conjunto de testes de largo espectro para a técnica da modelização híbrida desenvolvida.

Referências

- Beccari, M., Majone, M., Massanisso, P. e Ramadori, R. (1998), 'A bulking sludge with high storage response selected under intermittent feeding', *Water Research*, **32**(11), 3403–3413.
- Beun, J. J., Dircks, K., Van Loosdrecht, M. C. M. e Heijnen, J. J. (2002), 'Poly-*beta*-hydroxybutyrate metabolism in dynamically fed mixed microbial cultures', *Water Research*, **36**(5), 1167–1180.
- Beun, J. J., Paletta, F., Van Loosdrecht, M. C. M. e Heijnen, J. J. (2000), 'Stoichiometry and kinetics of poly-*beta*-hydroxybutyrate metabolism in aerobic, slow growing, activated sludge cultures', *Biotechnology and Bioengineering*, **67**(4), 379–389.
- Carucci, A., Dionisi, D., Majone, M., Rolle, E. e Smurra, P. (2001), 'Aerobic storage by activated sludge on real wastewater', *Water Research*, **35**(16), 3833–3844.
- Daigger, G. T. e Grady, C. P. L. (1982), 'An assessment of the role of physiological adaptation in the transient-response of bacterial cultures', *Biotechnology and Bioengineering*, **24**(6), 1427–1444.
- Ferreira, E. (1995), *Identificação e Controlo Adaptativo de Processos Biotecnológicos*, Tese de Doutoramento, Faculdade de Engenharia da Universidade do Porto.
- Gujer, W., Henze, M., Mino, T. e van Loosdrecht, M. (1999), 'Activated sludge model no. 3', *Water Science and Technology*, **39**(1), 183–193.
- Henze, M., Gujer, W., Mino, T., Matsuo, T., Wentzel, M. C., Marais, G. V. R. e Van Loosdrecht, M. C. M. (1999), 'Activated sludge model no.2d, asm2d', *Water Science and Technology*, **39**(1), 165–182.
- Henze, M., Gujer, W., Mino, T. e van Loosdrecht, M. E. (Editores) (2000), *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3; Scientific and Technical Report 9*, IWA Publishing, London.

- Krishna, C. e Van Loosdrecht, M. C. M. (1999a), 'Effect of temperature on storage polymers and settleability of activated sludge', *Water Research*, **33**(10), 2374–2382.
- Krishna, C. e Van Loosdrecht, M. C. M. (1999b), 'Substrate flux into storage and growth in relation to activated sludge modeling', *Water Research*, **33**(14), 3149–3161.
- Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., Zipursky, L. e Darnell, J. (2000), *Molecular Cell Biology*, W.H. Freeman.
- Oliveira, F. M. (1997), *Monitorização e Controlo de Fermentadores: Aplicação ao Fermento de Padeiro*, Tese de Doutoramento, Faculdade de Engenharia da Universidade do Porto.
- Park, S. e Ramirez, W. F. (1988), 'Optimal production of secreted protein in fed-batch reactors', *AIChE Journal*, **34**(9), 1550–1558.
- Park, S. e Ramirez, W. F. (1989), 'Dynamics of foreign protein secretion from *Saccharomyces-cerevisiae*', *Biotechnology and Bioengineering*, **33**(3), 272–281.
- Pomerleau, Y. e Perrier, M. (1990), 'Estimation of multiple specific growth-rates in bioprocesses', *AIChE Journal*, **36**(2), 207–215.
- Reed, G. e Pepler, H. (1973), *Yeast Technology*, The AVI Publishing Company, Inc., Connecticut.
- Serafim, L. S., Lemos, P. C., Oliveira, R. e Reis, M. A. M. (2004), 'Optimization of polyhydroxybutyrate production by mixed cultures submitted to aerobic dynamic feeding conditions', *Biotechnology and Bioengineering*, **87**(2), 145–160.
- Sonnleitner, B. e Kappeli, O. (1986), 'Growth of *Saccharomyces-cerevisiae* is controlled by its limited respiratory capacity formulation and verification of a hypothesis', *Biotechnology and Bioengineering*, **28**(6), 927–937.
- Sá-Correia, I., Moreira, L. e Fialho, A. (2003), 'Engenharia genética', em N. Lima e M. Mota (Editores), *Biotecnologia: Fundamentos e Aplicações*, LIDEL - Edições Técnicas, Lda., Páginas 125 – 161.
- Third, K. A., Newland, M. e Cord-Ruwisch, R. (2003), 'The effect of dissolved oxygen on phb accumulation in activated sludge cultures', *Biotechnology and Bioengineering*, **82**(2), 238–250.
- van Aalst-van Leeuwen, M. A., Pot, M. A., van Loosdrecht, M. C. M. e Heijnen, J. J. (1997), 'Kinetic modeling of poly(beta-hydroxybutyrate) production and consumption

- by *Paracoccus pantotrophus* under dynamic substrate supply', *Biotechnology and Bioengineering*, **55**(5), 773–782.
- van Loosdrecht, M. C. M. e Heijnen, J. J. (2002), 'Modelling of activated sludge processes with structured biomass', *Water Science and Technology*, **45**(6), 13–23.
- Walsh, G. (2002), *Proteins: Biochemistry and Biotechnology*, John Wiley Sons.

Capítulo 4

Modelização Híbrida de Processos (Bio)químicos: Definição de Uma Estratégia

Conteúdo do Capítulo

Processos biológicos que envolvem culturas de células são usualmente muito difíceis de modelizar essencialmente devido à complexidade dos fenómenos intracelulares assim como da heterogeneidade morfológica das populações de células. Quando os processos são complexos e pouco conhecidos do ponto de vista mecanístico, a modelização híbrida pode ser vantajosa com vantagem porque a exactidão do modelo pode aumentar com a incorporação de fontes de conhecimento alternativas e complementares. Neste capítulo é proposto um modelo híbrido dinâmico dum biorreactor que combina modelos mecanísticos com modelos não paramétricos: o biorreactor é descrito por um conjunto de equações de balanço material e a população celular é representada por uma mistura ajustável das representações mecanísticas e não paramétricas. Condições de estabilidade entrada limitada saída limitada (*Bounded Input Bounded Output* - BIBO) são derivadas para o modelo híbrido dinâmico geral. Duas estratégias para a identificação dos componentes não paramétricos são comparadas. As equações de sensibilidade são derivadas permitindo o cálculo analítico da matriz Jacobiana. A aplicação da teoria é ilustrada com dois casos de simulação.

4.1 Introdução

O desenvolvimento de estratégias avançadas baseadas em modelos para monitorização, optimização e controlo de biorreactores são por vezes condicionadas pela dificuldade do desenvolvimento dum bom modelo do processo. Frequentemente, os modelos de bioprocessos são excessivamente simplificados no que diz respeito à descrição do subsistema 'população celular' principalmente porque o desenvolvimento de modelos detalhados de células representam um custo proibitivo para este tipo de desenvolvimento.

Os modelos celulares podem ser classificados como estruturados/não-estruturados e segregados/não-segregados dependendo se os modelos entram em consideração ou não com a estrutura intracelular e com a heterogeneidade da população celular, respectivamente. Na maior parte das vezes, devido à falta de conhecimento básico ou devido à sua profunda complexidade, só os modelos não-segregados e não-estruturados são passíveis de ser utilizados.

Embora em alguns processos estacionários modelos simples são capazes de produzir resultados com exactidão suficiente, em processos semicontínuos altamente dinâmicos, negligenciar a estrutura intracelular e a heterogeneidade das células resultará quase invariavelmente num modelo com capacidades de estimação ou predição insuficientes.

Daí que, a modelização híbrida seja reconhecida como uma alternativa aos modelos mecanísticos para a análise de bioprocessos com custos de desenvolvimento inferiores (Schubert et al., 1994b,a; Preusting et al., 1996; Simutis et al., 1997; van Can et al., 1998; Peres et al., 2001). O princípio basilar da concepção deste tipo de modelos é o de não considerar o conhecimento mecanístico apriorístico como a única fonte de conhecimento relevante, considerando outras fontes de conhecimento - heurísticas ou informação escondida em bases de dados - como recursos complementares importantes, não alternativos, para o desenvolvimento de modelos. De facto, para sistemas complexos, para os quais há falta de conhecimento mecanístico, este ponto é particularmente importante.

Uma vez que os modelos híbridos integram mais conhecimento que os modelos clássicos baseados em princípios fundamentais, poder-se-á obter modelos mais exactos com menos experiências e portanto com custos de desenvolvimento inferiores.

A aplicação da modelização híbrida a reactores químicos e bioquímicos tem sido demonstrada em diversos trabalhos. A estrutura híbrida amplamente mais utilizada é baseada nas equações de balanço de massa, como na abordagem tradicional baseada em princípios fundamentais, no entanto as cinéticas são modelizadas por redes neuronais

artificiais (Psichogios e Ungar, 1992; Thompson e Kramer, 1994; Montague e Morris, 1994; Feyerherm et al., 1997; van Can et al., 1998, 1999; Chen et al., 2000).

Concretamente, as redes de perceptrão de camada múltipla (*Multiple Layer Perceptron* - MLP) e as redes de funções de base radial (*Radial Basis Function* - RBF) são as mais utilizadas por duas razões:

1. porque foi provado que estas redes são funções não lineares de aproximação universais;
2. a sua aplicação não requer o conhecimento acerca da estrutura do sistema que se pretende modelizar.

Estes dois argumentos motivaram a sua aplicação para modelizar os mecanismos muito complexos associados com o crescimento celular e biocatálise (Montague e Morris, 1994). A combinação de ANNs com equações de balanço material e de energia em paralelo (Thompson e Kramer, 1994) ou em série (Psichogios e Ungar, 1992) constitui-se em sistemas não lineares dinâmicos descritos por um conjunto de equações diferenciais ordinárias (*Ordinary Differential Equations* - ODEs).

Na perspectiva da identificação e análise de sistemas dinâmicos existem muitas questões teóricas tais como a identifiabilidade e a estabilidade que não estão bem estudadas mesmo para as estruturas híbridas mais simples ANN/equações de balanço. De facto, a maior parte dos estudos referidos são eminentemente orientados a um problema específico.

Neste capítulo é proposta uma estrutura de um modelo híbrido para um reactor que combina princípios fundamentais com modelos não paramétricos: o sistema biorreactor é descrito por um conjunto de equações de balanço material e o sistema população celular é representado por uma mistura ajustável de representações não paramétricas e mecanísticas. Além da formulação matemática desta estrutura híbrida, válida para uma larga gama de problemas, o sistema resultante é analisado numa perspectiva de engenharia de sistemas particularmente no que se refere à identificação dos subsistemas embutidos e sua estabilidade.

Todavia, um problema clássico desta técnica híbrida é a desobediência das restrições físicas do processo tais como as leis de conservação ou princípios termodinâmicos ou mesmo a positividade dos valores das concentrações (Thompson e Kramer, 1994; Karama et al., 2001). Por conseguinte, derivaram-se as condições de estabilidade entrada limitada saída limitada (BIBO) para o modelo híbrido geral garantindo valores das

concentrações estritamente positivos assim como limites superiores para os valores das concentrações. São discutidas duas estratégias para a identificação de parâmetros e é apresentada uma condição de identifiabilidade. Derivaram-se, também, para o modelo híbrido geral proposto neste trabalho, as equações das sensibilidades geralmente utilizadas para o cálculo das sensibilidades entrada/saída e parâmetro/saída dos modelos híbridos ANN/equações de balanço material (Psichogios e Ungar, 1992; Schubert et al., 1994a).

A aplicação da teoria é ilustrada com dois casos de estudo de simulação de processos de fermentação nomeadamente, a produção de proteína recombinante em modo semicontínuo e a produção de fermento de padeiro.

4.2 Modelo Híbrido Geral

A dinâmica dum biorreactor agitado pode ser descrita por um conjunto de equações de balanço material para um conjunto de componentes de interesse na fase líquida. Considerando que a fase líquida é perfeitamente agitada, as equações de balanço material podem ser representadas matricialmente em espaço de estado da seguinte forma:

$$\frac{d\mathbf{c}}{dt} = \mathbf{v}(\mathbf{c}) - D\mathbf{c} + \mathbf{u} \quad (4.1)$$

sendo \mathbf{c} um vector de n concentrações (o vector de estado), $\mathbf{v}(\mathbf{c})$ um vector de taxas volumétricas de reacção, D é a taxa de diluição e \mathbf{u} um vector de taxas volumétricas de entrada (taxas volumétricas de alimentação de nutrientes mais taxas volumétricas de transferência de massa gás-líquido).

A equação (4.1) é válida para os diferentes modos de operação de biorreactores:

- para a operação descontínua, o que corresponde a fazer $D = 0$;
- para a operação semicontínua, o que corresponde a

$$D = F/V, \quad dV/dt = F$$

sendo V o volume líquido e F a taxa de alimentação de entrada;

- para a operação contínua transitória, o que corresponde a ter

$$D = F_{in}/V = F_{out}/V, \quad dV/dt = F_{in} - F_{out} = 0$$

onde os subscritos 'in' e 'out' significam corrente de entrada e corrente de saída, respectivamente;

- e ainda para a operação contínua de estado estacionário onde

$$D = F_{in}/V = F_{out}/V \text{ e } d\mathbf{c}/dt = 0.$$

Com efeito, as equações de balanço material constituem o conhecimento mecanístico base que pode ser aplicado na maioria dos casos com elevados níveis de confiança. O mesmo não se pode dizer em relação às cinéticas $v(\mathbf{c})$. Bastin e Dochain (1990) propuseram uma metodologia e derivaram uma série de algoritmos de controlo seguindo a filosofia de 'modelização cinética mínima'. A ideia foi a de separar um termo cinético conhecido dum termo cinético desconhecido e depois elaborar algoritmos de estimação e controlo que não requerem o conhecimento do termo desconhecido.

Em consonância com este princípio, as taxas de reacção foram formuladas matematicamente pela equação:

$$v(\mathbf{c}) = \mathbf{KH}(\mathbf{c})\rho(\mathbf{c}) \quad (4.2)$$

sendo \mathbf{K} uma matriz de coeficientes de rendimento de dimensão $n \times r$, $\mathbf{H}(\mathbf{c})$ uma matriz de expressões cinéticas conhecidas de dimensão $r \times r$ e $\rho(\mathbf{c})$ um vector de r funções cinéticas desconhecidas. Neste estudo adoptou-se a Equação (4.2) para definir também o termo da reacção. Ao fazê-lo não se tem a intenção de adoptar a filosofia da 'modelização cinética mínima' mas sim separar um termo mecanístico conhecido doutro termo que pode ser modelizado com técnicas alternativas.

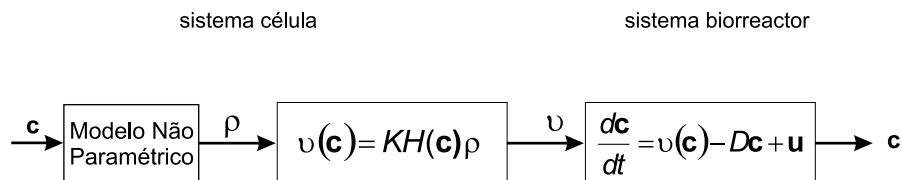


Figura 4.1: Estrutura do modelo híbrido geral para biorreactores agitados

Como já foi discutido atrás, a principal questão na concepção da modelização híbrida é o de dever permitir incorporar diferentes fontes de conhecimento num só modelo. Conjugando esta questão com o reconhecimento da generalidade da Equação (4.2), é proposta uma estrutura híbrida geral representada esquematicamente na Figura 4.1. Esta estrutura híbrida pode ser expressa matematicamente pelas duas equações seguintes:

$$\frac{d\mathbf{c}}{dt} = \mathbf{KH}(\mathbf{c})\rho - D\mathbf{c} + \mathbf{u} \quad (4.3)$$

$$\rho = \mathbf{N}(\mathbf{c}, \mathbf{W}) \quad (4.4)$$

Concretamente, esta estrutura híbrida deve obedecer aos seguintes princípios de concepção: todo o conhecimento baseado em princípios fundamentais deve ser incorporado na Equação (4.3) enquanto que todas as outras fontes de conhecimento devem ser incorporadas na Equação (4.4). A Equação (4.4) estabelece que o termo ρ seja calculado por uma técnica não paramétrica, por exemplo uma rede MLP. No entanto, a técnica não paramétrica refere-se a sistemas conexionistas em geral: não só a redes neuronais artificiais mas também a arquitecturas modulares, a sistemas difusos ou modelos estatísticos podem ser considerados na definição da função da Equação (4.4). Com este formalismo matemático, o conhecimento mecanístico é usado preferencialmente, enquanto que os outros tipos de conhecimento são activados no modelo através da Equação (4.4).

Podem-se realçar três propriedades importantes do sistema (4.3)-(4.4):

1. A representação do termo cinético através da Equação (4.2) é bastante genérica tanto para reacções de catálise químicas como biológicas (p.e., Bastin e Dochain, 1990; Dochain et al., 1992).
2. A organização introduzida por esta expressão permite o uso de outras técnicas de modelização para calcular ρ . Daí que, se possa utilizar em vez de uma única rede neuronal, m redes neuronais, um sistema difuso ou uma série de combinações destas estruturas ou doutras.
3. Desde que todas as funções na matriz $\mathbf{N}(\mathbf{c}, \mathbf{W})$ sejam contínuas, diferenciáveis e limitadas, a análise de estabilidade entrada limitada saída limitada (BIBO) pode ser efectuada e, ainda mais importante, as sensibilidades dos parâmetros podem ser calculadas. Esta questão vai ser discutida em detalhe na Secção 4.3.

De seguida descrevem-se as duas principais redes neuronais artificiais que se utilizam neste trabalho para representar ρ na Equação (4.4).

4.2.1 Redes de Percepção de Camada Múltipla (MLP)

As redes de Percepção de Camada Múltipla são uma classe importante de redes neuronais artificiais. Tipicamente, as redes MLP consistem numa camada de entrada, numa ou mais camadas internas e numa camada de saída como apresentado esquematicamente na Figura 4.2. A organização dos nodos em camadas e as ligações entre as camadas é que dão a este tipo de redes a capacidade de mapear sistemas de entrada/saída.

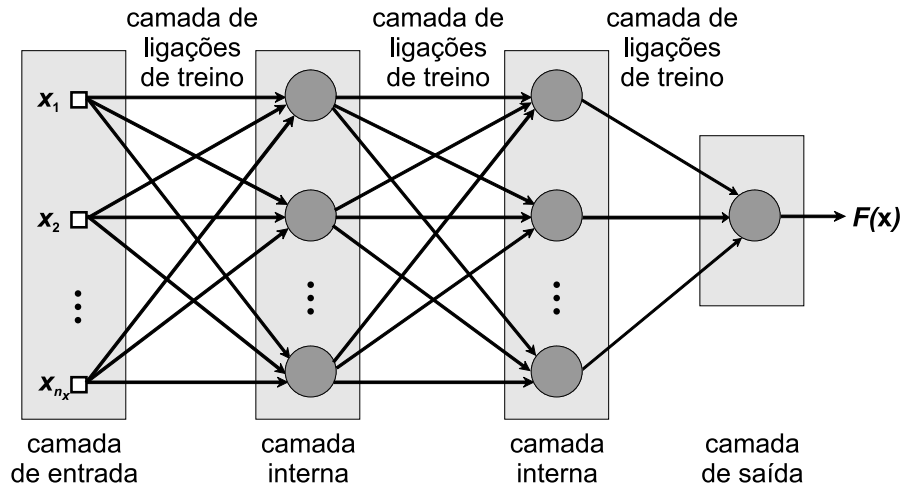


Figura 4.2: Arquitectura duma rede de perceptron de camada múltipla com duas camadas internas e um neurónio de saída

Um qualquer número de entradas pode ser usado para gerar um qualquer número de saídas. Por isso, o que define a arquitectura da rede é o número de nodos em cada camada e o número de camadas internas. Por exemplo, $\{3, 5, 4\}$ define uma rede MLP com 3 entradas, 5 nodos na camada interna e 4 saídas. O modelo de cada nodo da rede inclui uma função de activação tipicamente não linear. Esta função de activação tem por norma a característica importante da não linearidade ser 'suave', isto é, diferenciável em todo o seu domínio (Haykin, 1999). As funções de activação mais utilizadas são a função tangente hiperbólica dada pela seguinte expressão:

$$s(x) = \tanh(x) = \frac{\exp(x) - \exp(-x)}{\exp(x) + \exp(-x)} \quad (4.5)$$

e a função sigmóide:

$$s(x) = \frac{1}{1 + \exp(-x)}. \quad (4.6)$$

A saída $F(\mathbf{x})$ da rede MLP representada na Figura 4.2 pode ser definida da seguinte forma:

$$F(\mathbf{x}) = s(\mathbf{W}_2 s(\mathbf{W}_1 s(\mathbf{S}) + \mathbf{B}_1) + \mathbf{B}_2) \quad (4.7)$$

onde \mathbf{W}_1 , \mathbf{B}_1 são as matrizes de parâmetros associados às ligações entre os nodos da camada de entrada e da camada interna, \mathbf{W}_2 , \mathbf{B}_2 são as matrizes de parâmetros associados às ligações entre os nodos da camada interna e a camada de saída da rede neuronal MLP, e s a função sigmóide. O vector de parâmetros \mathbf{W} da Figura 4.2 é a

representação vectorial das matrizes \mathbf{W}_1 , \mathbf{B}_1 , \mathbf{W}_2 , \mathbf{B}_2 .

As redes MLP têm sido aplicadas com sucesso para resolver alguns problemas complexos, nomeadamente no campo da biotecnologia, treinando-as em modo de supervisão com um algoritmo muito popular denominado algoritmo da retropropagação do erro (Rumelhart et al., 1986). O desenvolvimento deste algoritmo representa um marco histórico na medida em que é uma técnica computacionalmente eficiente para o treino de redes MLP (Haykin, 1999).

4.2.2 Redes de Funções de Base Radial (RBF)

Uma rede de Funções de Base Radial envolve na sua forma básica três camadas como representado no esquema da Figura 4.3. A camada de entrada composta pelos nodos de entrada de igual número à dimensão n_x do vector \mathbf{x} (i.e., o número das variáveis independentes do problema), a camada interna composta por m_1 unidades não lineares onde cada unidade está ligada directamente a todos os nodos de entrada. As funções de activação da camada interna são as funções de base radial φ . A camada de saída consiste numa única saída linear ligada completamente à camada interna, i.e., a saída da rede $F(\mathbf{x})$ é a soma linear pesada das saídas das unidades internas:

$$F(\mathbf{x}) = \sum_{i=1}^{m_1} w_i \varphi_i(\mathbf{x}) \quad (4.8)$$

em que

$$\varphi_i(\mathbf{x}) = G(\|\mathbf{x} - \mathbf{t}_i\|), \quad i = 1, 2, \dots, m_1 \quad (4.9)$$

e em que o conjunto dos centros $\{\mathbf{t}_i \mid i = 1, 2, \dots, m_1\}$ têm de ser determinados. G é uma função definida para argumentos positivos.

Concretamente, as funções de base radial φ podem ser da forma (Haykin, 1999):

1. Multiquadratica:

$$\varphi_i(\|\mathbf{x} - \mathbf{t}_i\|) = (\|\mathbf{x} - \mathbf{t}_i\|^2 + c^2)^{1/2} \text{ para algum } c \geq 0 \text{ e } r \in \mathbb{R} \quad (4.10)$$

2. Multiquadratica inversa:

$$\varphi_i(\|\mathbf{x} - \mathbf{t}_i\|) = \frac{1}{(\|\mathbf{x} - \mathbf{t}_i\|^2 + c^2)^{1/2}} \text{ para algum } c \geq 0 \text{ e } r \in \mathbb{R} \quad (4.11)$$

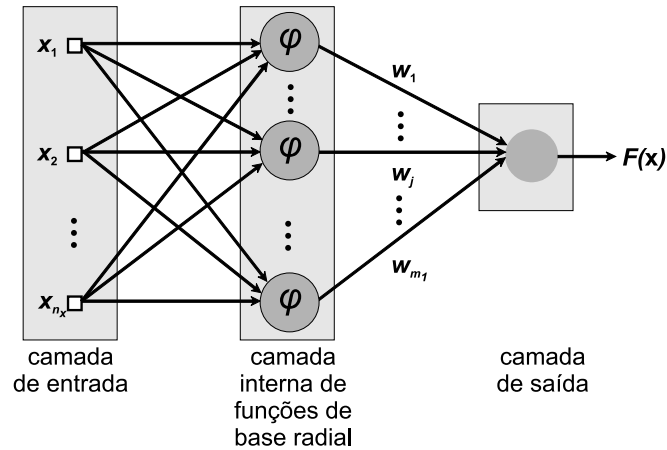


Figura 4.3: Arquitectura duma rede de funções de base radial

3. Funções Gaussianas:

$$\varphi_i(\|\mathbf{x} - \mathbf{t}_i\|) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{t}_i\|^2}{2\sigma^2}\right) \text{ para algum } \sigma \geq 0 \text{ e } \mathbf{t}_i \in \mathbb{R} \quad (4.12)$$

O número de nodos da camada interna é duma maneira geral muito menor que o número de pontos medidos. Os pesos lineares associados à camada de saída, a posição dos centros das funções de base radial e a matriz G associadas à camada interna são tudo parâmetros desconhecidos que serão determinados através do treino da rede.

As redes RBF e as redes MLP são ambas exemplos de redes não lineares em camadas alimentadas para a frente e são ambas aproximadores universais. A principal diferença entre estas duas redes é que a rede MLP constrói aproximações globais do mapeamento não linear entrada/saída enquanto que as redes RBF constroem aproximações locais do mesmo mapeamento (Haykin, 1999).

4.3 Derivação das condições de estabilidade BIBO

É reconhecido que a modelização com redes neuronais artificiais, e em geral da modelização não paramétrica, têm a desvantagem de violarem certas restrições físicas, como por exemplo, as fracções molares e mássicas terem de somar um e os valores das concentrações terem de ser positivos (Thompson e Kramer, 1994; Karama et al., 2001; Feyo de Azevedo et al., 1997). Com efeito, na modelização híbrida existe o mesmo problema embora atenuado de alguma forma devido à inclusão de algum conhecimento físico apriorístico na estrutura do modelo (Thompson e Kramer, 1994).

Daí que, no contexto deste trabalho seja importante analisar as condições sob as quais as saídas do modelo híbrido dinâmico geral são limitadas e, derivar se possível restrições na estrutura do modelo ou do processo que garantam a positividade das concentrações. Concretamente, isto pode ser realizável através do estudo da análise de estabilidade BIBO.

Bastin e Dochain (1990) efectuaram um estudo da análise de estabilidade BIBO para um sistema dinâmico equivalente aos sistemas (4.1) e (4.3). As condições sob as quais as concentrações são estritamente positivas e limitadas superiormente podem ser resumidas assim:

- C1. A taxa de diluição D é limitada inferiormente;
- C2. As correntes de entrada \mathbf{u} são limitadas superiormente;
- C3. Cada reacção r pode ocorrer se e só se todos os substratos estiverem presentes no meio e cada reacção envolver pelo menos um substrato que não é um catalisador (i.e., as células).

As condições C1 e C2 são condições relacionadas com a operação. No entanto, a condição C3 impõe algumas restrições na definição da Equação (4.4). A condição C3 é verificada se os elementos da diagonal da matriz \mathbf{H} (notar que a matriz \mathbf{H} é por definição diagonal e os elementos da diagonal representam funções cinéticas conhecidas) forem iguais ao produto das concentrações $c_{k,j}$ de todos os substratos k que intervêm na reacção j (Equação (4.13)).

$$\mathbf{H}(\mathbf{c}) = \text{diag} \left\{ \prod_k c_{k,j} \right\} \quad j = 1, \dots, r, \quad i = 1, \dots, n \quad (4.13)$$

Daí que, com esta definição as taxas volumétricas de reacção $v_j(\mathbf{c}) = h_j(\mathbf{c})\rho_j(\mathbf{c})$ sejam zero sempre que um reagente da reacção se esgotar e assim a condição C3 é verificada.

4.4 Identificação de Parâmetros

A equação (4.4) estabelece uma relação não linear paramétrica ou semiparamétrica entre ρ e \mathbf{c} onde um conjunto de parâmetros \mathbf{W} de dimensão n_w estão envolvidos. Por sua vez, estes parâmetros serão identificados através de dados. Como é sabido, as

cinéticas do processo não podem ser medidas directamente, só as concentrações podem ser medidas directamente. Daí que, as taxas de reacção ρ estejam relacionadas com as quantidades \mathbf{c} que são medidas pela Equação (4.3). Contudo, na prática apenas uma partição com r equações são necessárias e sendo assim ρ é calculado pela expressão

$$\rho = [\mathbf{K}_a \mathbf{H}(\mathbf{c})]^{-1} \left(-\frac{d\mathbf{c}_a}{dt} + D\mathbf{c}_a - \mathbf{u}_a \right) \quad (4.14)$$

em que o índice a denota a partição com r variáveis de estado. Da Equação (4.14) resulta a seguinte condição de identifiabilidade do modelo híbrido:

- C4. O modelo híbrido geral (4.3)-(4.4) é identificável se e só se existir uma partição a de r variáveis de estado tal que a matriz correspondente $\mathbf{K}_a \mathbf{H}(\mathbf{c})$ de dimensão $r \times r$ é não-singular.

Independentemente do tipo de relação definida na Equação (4.4) o objectivo do procedimento de identificação é o de obter o vector dos parâmetros \mathbf{W} que minimiza o desvio entre as saídas do modelo e do processo. De facto, duas estratégias são possíveis para esta identificação. Estas serão descritas nas Subsecções 4.4.1 e 4.4.2 seguintes.

4.4.1 Estratégia I: erro dos mínimos quadrados em ρ

A estratégia I consiste na minimização directa, geralmente no sentido dos mínimos quadrados, dos erros entre as taxas de reacção 'medidas' e modelizadas, assumindo erros de medida Gaussianos.

Para um conjunto de P padrões medidos $\{\mathbf{c}_t, \rho_t\}$ a função do erro dos mínimos quadrados é definida da seguinte forma:

$$\arg \min_{\mathbf{W}} \left\{ J = \frac{1}{P} \sum_{t=1}^P [\rho_t - \mathbf{N}(\mathbf{c}_t, \mathbf{W})]^T \Sigma [\rho_t - \mathbf{N}(\mathbf{c}_t, \mathbf{W})] \right\} \quad (4.15)$$

em que Σ é a matriz das covariâncias do ruído nas medidas. A identificação de \mathbf{W} pode ser realizada através de algoritmos de treino de redes neuronais artificiais tais como o algoritmo de retropropagação com validação cruzada ou outros (Leonard et al., 1992; Pollard et al., 1992; Qin et al., 1992).

Como mencionado previamente, ρ_t não é medido directamente mas pode ser estimado a partir das concentrações \mathbf{c} medidas usando a Equação (4.14). Por sua vez, a solução da Equação (4.14) exige uma aproximação para as derivadas das concentrações em ordem

ao tempo incluídas na equação. Uma hipótese é utilizar, após a remoção do ruído, algoritmos numéricos comuns de diferenciação.

Alternativamente pode-se usar um filtro não linear para a estimação das taxas de reacção desconhecidas. Concretamente, para o modelo híbrido dinâmico (4.3)-(4.4), pode ser derivado um estimador baseado num observador (*observer-based estimator* - OBE) (Bastin e Dochain, 1990) assumindo a seguinte forma geral:

$$\frac{d\hat{\mathbf{c}}_a}{dt} = \mathbf{K}_a \mathbf{H}(\mathbf{c}) \hat{\rho} - D \mathbf{c}_a + \mathbf{u}_a + \mathbf{\Omega}(\mathbf{c}_a - \hat{\mathbf{c}}_a) \quad (4.16)$$

$$\frac{d\hat{\rho}}{dt} = [\mathbf{K}_a \mathbf{H}(\mathbf{c})]^T \mathbf{\Gamma}(\mathbf{c}_a - \hat{\mathbf{c}}_a) \quad (4.17)$$

em que o símbolo $\hat{\cdot}$ significa quantidades estimadas e os símbolos $\mathbf{\Omega}$ e $\mathbf{\Gamma}$ são matrizes diagonais de parâmetros à disposição do utilizador para sintonizar a velocidade de convergência e as propriedades de seguimento do estimador. Com uma sintonização apropriada é possível obter respostas quase de segunda ordem com constante de tempo e coeficiente de amortecimento configuráveis, os quais estão restringidos pelo tempo de amostragem (Oliveira et al., 2002). A aplicação da estratégia I usando o estimador OBE foi exemplificada por Chen et al. (2000).

4.4.2 Estratégia II: erro dos mínimos quadrados em \mathbf{c}

Esta estratégia consiste em minimizar o desvio entre os valores medidos das concentrações e os obtidos pelo modelo. Para um conjunto P de padrões medidos $\{\mathbf{c}_{m,t}, \mathbf{F}_t, \mathbf{u}_t\}$ o critério do mínimos quadrados define-se da seguinte forma:

$$\arg \min_{\mathbf{W}} \left\{ J = \frac{1}{P} \sum_{t=1}^P [\mathbf{c}_{m,t} - \mathbf{c}_t]^T \Sigma [\mathbf{c}_{m,t} - \mathbf{c}_t] \right\} \quad (4.18)$$

sendo $\mathbf{c}_{m,t}$ os valores medidos das concentrações. Repare-se que os valores das concentrações estimados pelo modelo híbrido \mathbf{c}_t são variáveis dinâmicas dependentes de \mathbf{W} através das Equações (4.3) e (4.4). Assim a prossecução da estratégia II requer que as equações do modelo híbrido (4.3)-(4.4) sejam integradas numericamente entre os valores medidos em oposição com a estratégia I que requer diferenciação numérica. Estratégias de optimização estocásticas ou baseadas em gradientes podem ser empregues para resolver a optimização (4.18) que pode ser vista como um problema de programação semi-infinito (Thompson e Kramer, 1994).

A avaliação do erro de modelização através da Equação (4.18) pode consumir muito

tempo quando são empregues redes de grande dimensão e quando está disponível uma grande quantidade de valores medidos. Contudo, o treino de ANNs pode ser melhorado empregando métodos de optimização de gradientes de primeira ordem com linha de procura baseada em gradientes conjugados que no caso das ANNs podem ser calculados com o já muito conhecido algoritmo de retropropagação do erro (Leonard e Kramer, 1990). Psychogios e Ungar (1992) sugeriram o uso do método das sensibilidades para o cálculo dos gradientes $\partial J / \partial \mathbf{W}$ num modelo híbrido 'equações de balanço de massa/ANN' e empregaram programação não linear com a informação dos gradientes obtidos pelas equações das sensibilidades.

Concretamente, para o caso do modelo híbrido geral (4.3)-(4.4), as equações das sensibilidades podem ser derivadas desde que as funções $\mathbf{N}(\mathbf{c}, \mathbf{W})$ sejam contínuas e diferenciáveis. A diferenciação de J em ordem aos parâmetros \mathbf{W} resulta na equação seguinte:

$$\frac{\partial J}{\partial \mathbf{W}} = \sum_{t=1}^P \left(\frac{\partial J}{\partial \mathbf{c}} \right)_t \left(\frac{\partial \mathbf{c}}{\partial \mathbf{W}} \right)_t = -\frac{2}{P} \sum_{t=1}^P \mathbf{e}_t^T \left(\frac{\partial \mathbf{c}}{\partial \mathbf{W}} \right)_t \quad (4.19)$$

em que $\mathbf{e}_t = (\mathbf{c}_{m,t} - \mathbf{c}_t)$. A matriz $(\partial \mathbf{c} / \partial \mathbf{W})_t$ pode ser calculada através das equações das sensibilidades as quais podem ser obtidas por diferenciação das Equações (4.3) e (4.4) em ordem a \mathbf{W} . Sendo assim, após algumas manipulações matemáticas as seguintes equações são obtidas :

$$\frac{d}{dt} \left(\frac{\partial \mathbf{c}}{\partial \mathbf{W}} \right) = \mathbf{A} \frac{\partial \mathbf{c}}{\partial \mathbf{W}} + \mathbf{B} \quad (4.20)$$

em que,

$$\mathbf{A} = \mathbf{K} \mathbf{I}_r \rho^T \frac{\partial \mathbf{H}}{\partial \mathbf{c}} + \mathbf{K} \mathbf{H} \frac{\partial \rho}{\partial \mathbf{c}} - D \mathbf{I}_n \quad (4.21)$$

e

$$\mathbf{B} = \mathbf{K} \mathbf{H} \frac{\partial \rho}{\partial \mathbf{W}} \quad (4.22)$$

juntamente com as seguintes definições de matrizes:

$$\begin{aligned} \frac{\partial \mathbf{c}}{\partial \mathbf{W}} &= \left\{ \frac{\partial c_i}{\partial w_j} \right\} & \frac{\partial \mathbf{H}}{\partial \mathbf{c}} &= \left\{ \frac{\partial h_k}{\partial c_i} \right\} \\ \frac{\partial \rho}{\partial \mathbf{c}} &= \left\{ \frac{\partial N_k}{\partial c_i} \right\} & \frac{\partial \rho}{\partial \mathbf{W}} &= \left\{ \frac{\partial N_k}{\partial c_j} \right\} \end{aligned}$$

em que $i = 1, \dots, n$, $k = 1, \dots, r$ e $j = 1, \dots, nw$.

O conjunto de Equações (4.20) - (4.22) têm de ser integradas simultaneamente com as Equações (4.3) e (4.4). Como o valor inicial das variáveis de estado é independente dos parâmetros \mathbf{W} , as condições iniciais da Equação (4.20) são iguais a zero, isto é,

$$\left(\frac{\partial \mathbf{c}}{\partial \mathbf{W}} \right)_{t=0} = 0 \quad (4.23)$$

4.5 Caso de Estudo I: produção de proteína recombinante em modo semicontínuo

4.5.1 Descrição do processo

O modelo matemático descrito em Park e Ramirez (1988) e apresentado no Capítulo 3, Secção 3.1, para a produção de proteína recombinante em modo semicontínuo vai servir como exemplo para testar, por estudos de simulação, nas suas linhas gerais os métodos propostos.

As equações de balanço material (Equações (3.1) - (3.4)) tomam a seguinte forma na representação em espaço dos estados:

$$\frac{d}{dt} \begin{bmatrix} X \\ S \\ P_t \\ P_m \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -7.3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X & 0 & 0 \\ 0 & X & 0 \\ 0 & 0 & (P_t - P_m) \end{bmatrix} \times \begin{bmatrix} \mu(S) \\ f_p(S) \\ \Phi(S) \end{bmatrix} - D \begin{bmatrix} X \\ S - S_o \\ P_t \\ P_m \end{bmatrix} \quad (4.24)$$

em que X é a concentração de biomassa por unidade de volume do reactor; S é a concentração de glucose por unidade de volume do reactor; P_t a concentração total de proteína por unidade de volume do reactor; P_m a concentração de proteína excretada por unidade de volume de reactor, D a taxa de diluição ($D = F/V$ sendo F a taxa de alimentação e V o volume do meio dentro do biorreactor) e S_o a concentração de substrato na corrente de entrada. Uma vez que, o processo é operado em modo semicontínuo a equação de balanço material global $dV/dt = F = DV$ tem de ser integrada juntamente com a Equação (4.24).

As expressões cinéticas 'verdadeiras' são definidas pelas Equações (3.6) - (3.8) do Capítulo 3.

Foram efectuadas duas corridas que simularam 16 h de operação. Assumiram-se períodos de amostragem de 1 min para as medidas em linha de F e V (para ser preciso

V não foi medido mas integrado em diferido e fornecido como medida indirecta ao procedimento de identificação discutido em baixo) e períodos de amostragem de 15 min para as medidas em diferido das concentrações de X , S , P_t e P_m .

Cada partida ficou com o número total de $P = 960$ pontos medidos. Com o intuito de excitar o sistema e para obter variações largas nas concentrações de glucose, tendo em vista o estudo de identificação subsequente, a taxa de alimentação F é representada por uma função degrau na gama de 0.01 – 0.2 L/h controlada pela concentração de glucose e produzindo variações de glucose na gama de 0.1 – 10 g/L. Uma vez que as taxas cinéticas representadas pelas Equações (3.6) - (3.8) só dependem de S este procedimento vai restringir o domínio da experiência no passo da identificação. A concentração de glucose na corrente de entrada foi de $S_0 = 40$ g/L.

As condições iniciais das concentrações de biomassa e glucose foram escolhidas aleatoriamente a partir duma distribuição uniforme dentro dos intervalos 0 – 2 e 0 – 0.5 g/L, respectivamente. As concentrações iniciais para as proteínas total e segregada foram de $P_t(0) = 0$ e $P_m(0) = 0$, respectivamente. Às concentrações de X , S , P_t e P_m foi adicionado erro Gaussiano com desvio padrão de 0.25.

4.5.2 Estrutura do modelo híbrido

A estrutura do modelo híbrido foi derivada partindo do pressuposto que tanto as equações de balanço material (Equação (4.24)) como a estequiometria da reacção é conhecida. Daí que, a única parte do processo que é desconhecida do ponto de vista mecanístico sejam as taxas cinéticas expressas pelas Equações (3.6)- (3.8) do Capítulo 3.

Assim sendo, a estrutura do modelo híbrido geral (4.3) - (4.4) pode ser aplicada directamente fazendo

$$\mathbf{c} = \begin{bmatrix} X & S & P_t & P_m \end{bmatrix}^T$$

e

$$\mathbf{u} = \begin{bmatrix} 0 & DS_0 & 0 & 0 \end{bmatrix}^T.$$

A matriz das expressões cinéticas conhecidas é

$$\mathbf{H} = \text{diag} \left(\begin{bmatrix} X & X & (P_t - P_m) \end{bmatrix} \right)$$

e o vector das funções cinéticas desconhecidas é

$$\rho = \begin{bmatrix} \mu(S) & f_p(S) & \Phi(S) \end{bmatrix}^T.$$

As três taxas cinéticas desconhecidas da Equação (4.4) foram modelizadas com uma rede MLP standard com uma entrada (concentração de glucose), oito nodos internos e três saídas. Assim, o modelo híbrido consiste na Equação (4.24) mais a equação MLP adicional:

$$\begin{bmatrix} \mu & f_p & \Phi \end{bmatrix}^T = \text{diag} \left(\begin{bmatrix} \mu_{max} & f_{p,max} & \Phi_{max} \end{bmatrix} \right) \times s(\mathbf{W}_2 s(\mathbf{W}_1 s(\mathbf{S}) + \mathbf{B}_1) + \mathbf{B}_2) \quad (4.25)$$

em que \mathbf{W}_1 , \mathbf{B}_1 , \mathbf{W}_2 , \mathbf{B}_2 são as matrizes de parâmetros associados às ligações entre os nodos da rede neuronal MLP, e $s(x) = 1/(1 + e^{-x})$ a função sigmóide. O vector de parâmetros \mathbf{W} é a representação vectorial das matrizes \mathbf{W}_1 , \mathbf{B}_1 , \mathbf{W}_2 , \mathbf{B}_2 e compreende neste caso 42 parâmetros escalares.

4.5.3 Resultados da identificação

O primeiro estudo teve como objectivo identificar o vector de parâmetros \mathbf{W} utilizando a estratégia I. Contudo, foi impossível obter boas estimativas das cinéticas desconhecidas com o estimador OBE porque os dados tinham ruído e porque o período de amostragem era demasiado grande para capturar a dinâmica do processo, especialmente a dinâmica de P_t e de P_m .

Obtiveram-se os mesmos resultados insatisfatórios usando directamente a Equação (4.14) e empregando tanto ajustes por splines cúbicas, diferenciação directa de Euler ou diferenciação do ponto médio para a estimação de $d\mathbf{c}_a/dt$. Esta estimação, neste caso, corresponde à estimação de dX/dt , dS/dt , dP_t/dt e dP_m/dt .

Daí que, a amostragem em diferido e o comportamento altamente dinâmico imposto ao processo pelas alterações em degrau da taxa de alimentação F obsta a que se aplique a estratégia I.

No entanto, os resultados obtidos com a estratégia II foram bastante promissores. Foi utilizado um algoritmo de programação quadrática sequencial (*Sequential Quadratic Programming* - SQP) para resolver a optimização (4.18) utilizando os gradientes analíticos $\partial J/\partial \mathbf{W}$ da função objectivo fornecidos pelas equações das sensibilidades (4.20) - (4.22).

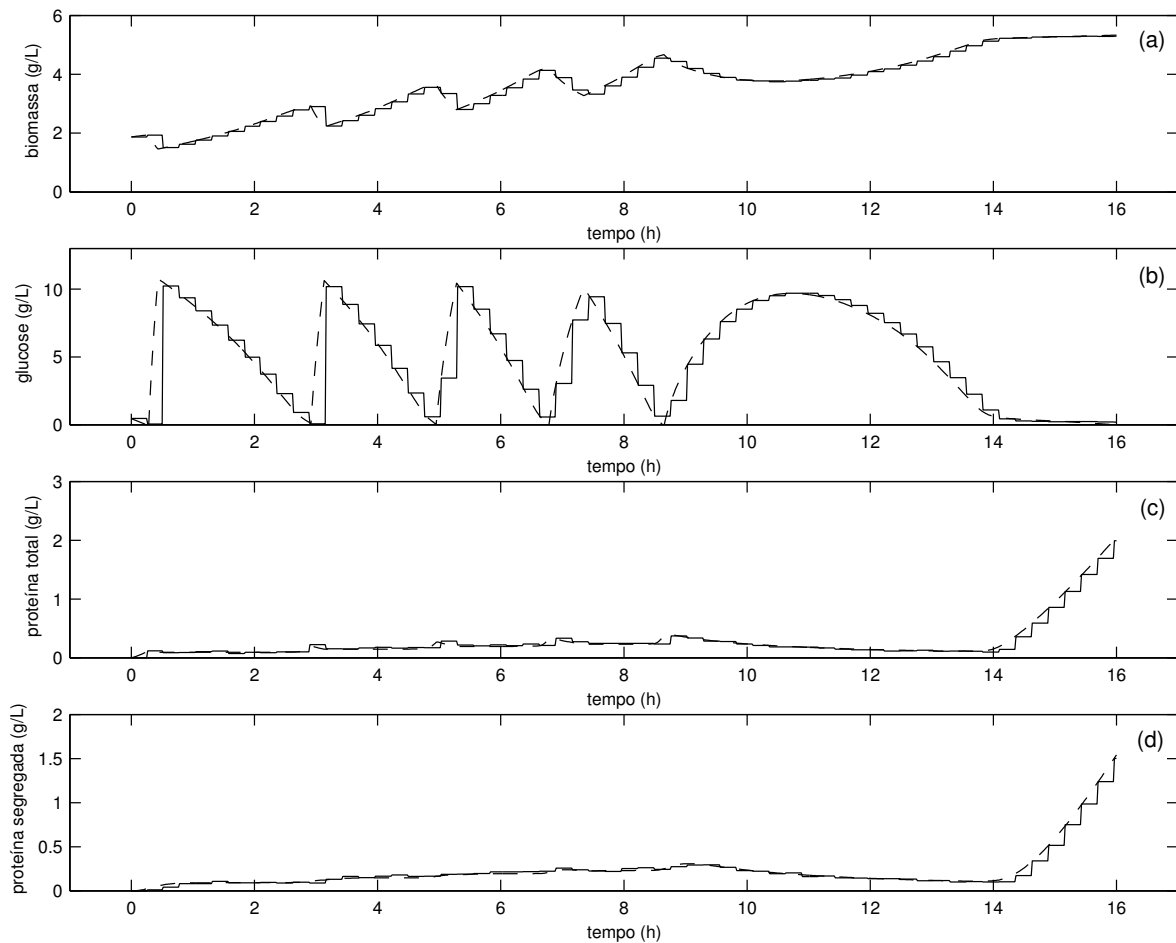


Figura 4.4: Resultados do modelo híbrido para a corrida de treino: (a) biomassa; (b) glucose; (c) proteína total; (d) proteína segregada. As linhas a cheio representam os valores 'medidos' e as linhas a tracejado representam as saídas do modelo híbrido

Neste estudo foi usada uma corrida do processo para a identificação e outra corrida de teste para fazer a validação do modelo. Os resultados obtidos com as corridas de treino e de teste apresentam-se nas Figuras 4.4 e 4.5, respectivamente.

O erro quadrático médio obtido para o conjunto de dados de teste foi de 4.7×10^{-3} (onde os valores das concentrações são escaladas pelos seus valores máximos). Embora o conjunto de dados usado para treino (ver Figura 4.4) difira bastante do conjunto de dados usado para teste (ver Figura 4.5), o modelo híbrido consegue descrever de modo exacto a dinâmica de todas as variáveis de estado do conjunto de dados de teste. Este resultado não é de todo inesperado uma vez que a gama dos valores da concentração de S é bastante extensa.

É de notar que a identificação se confina ao subsistema (4.4) que no caso presente só tem uma variável de entrada - S .

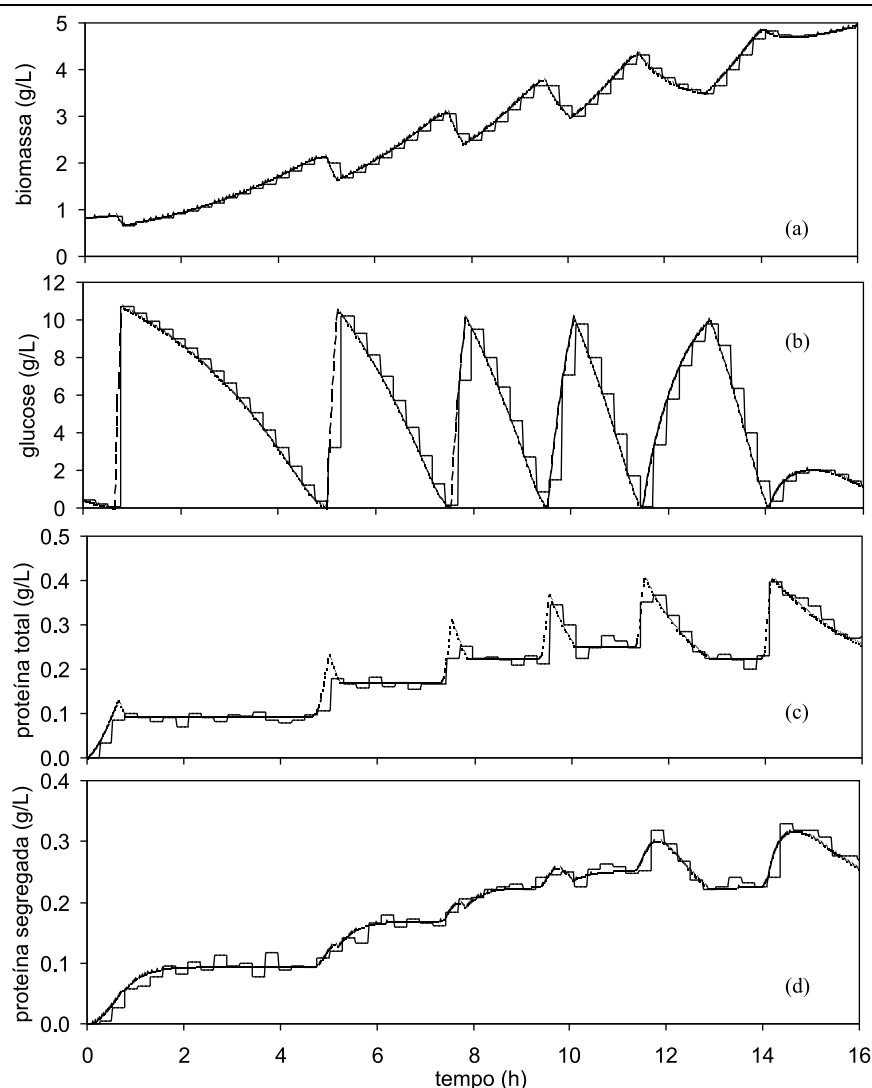


Figura 4.5: Resultados do modelo híbrido para a corrida de teste: (a) biomassa; (b) glucose; (c) proteína total; (d) proteína segregada. As linhas a cheio representam os valores 'medidos' e as linhas a tracejado representam as saídas do modelo híbrido

É de realçar também que a identificação foi bem sucedida utilizando a estratégia II - pelo menos parcialmente como se vai mostrar mais abaixo - apesar dos valores medidos serem esparsos e terem ruído e com a estratégia I falhou usando os mesmos dados. Contudo a estratégia I tem potencial no caso de se possuir medidas fiáveis em linha das variáveis de estado o que não é comum existir nas indústrias bioquímicas.

Mais interessante será analisar a identificação das taxas cinéticas desconhecidas. A Figura 4.6 mostra as funções cinéticas identificadas em conjunto com as curvas 'verdadeiras' ((3.6) - (3.8)) em função da concentração de glucose.

Verifica-se que as curvas 'verdadeiras' foram bem identificadas no que diz respeito à taxa específica de crescimento (μ) e à taxa específica de produção de proteína (f_p),

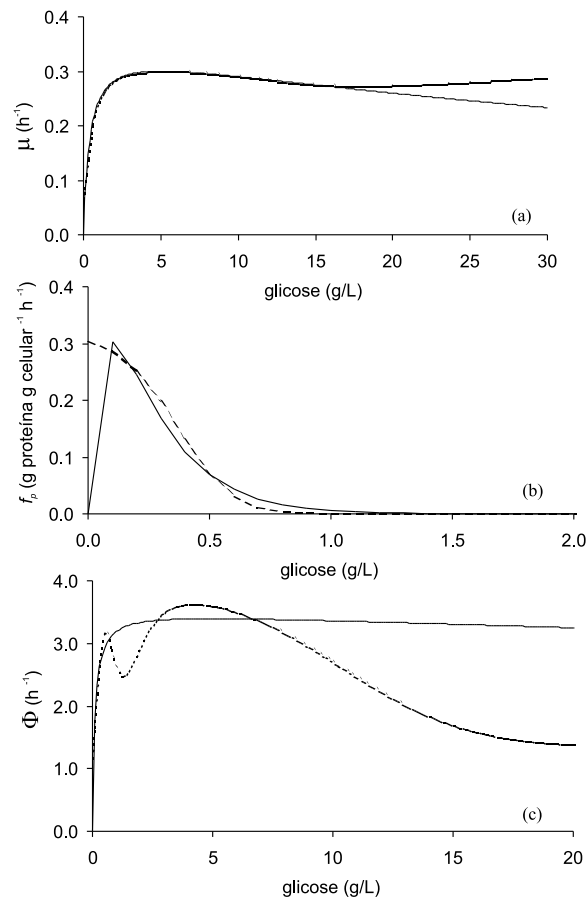


Figura 4.6: Resultados da modelização cinética: (a) taxa específica de crescimento; (b) taxa específica de produção de proteína; (c) taxa de secreção de proteína. As linhas a cheio representam as cinéticas verdadeiras e as linhas a tracejado representam as saídas do modelo híbrido

tendo-se no entanto, obtido melhores resultados para μ do que para f_p .

No caso da taxa específica de crescimento a exactidão da modelização degrada-se para valores de concentração de glucose superiores a 15 g/L. Isto pode ser explicado pelo facto de não haver medidas de concentrações de glucose nesta gama de valores como mostra a Figura 4.5.

Também para o caso da taxa específica de produção de proteína, os resultados da modelização degradam-se para valores de concentração de glucose baixos (Figura 4.6). Note-se que S foi controlado dentro da gama de 0.1 – 10 g/L e sendo assim poucos valores medidos de glucose existem abaixo de 0.1 g/L.

Estes resultados estão em consonância com os resultados obtidos noutros trabalhos referentes a identificação com redes neuronais: as redes são poderosas a interpolar mas pobres a extrapolar aliás como outras técnicas de modelização não paramétricas (Pollard

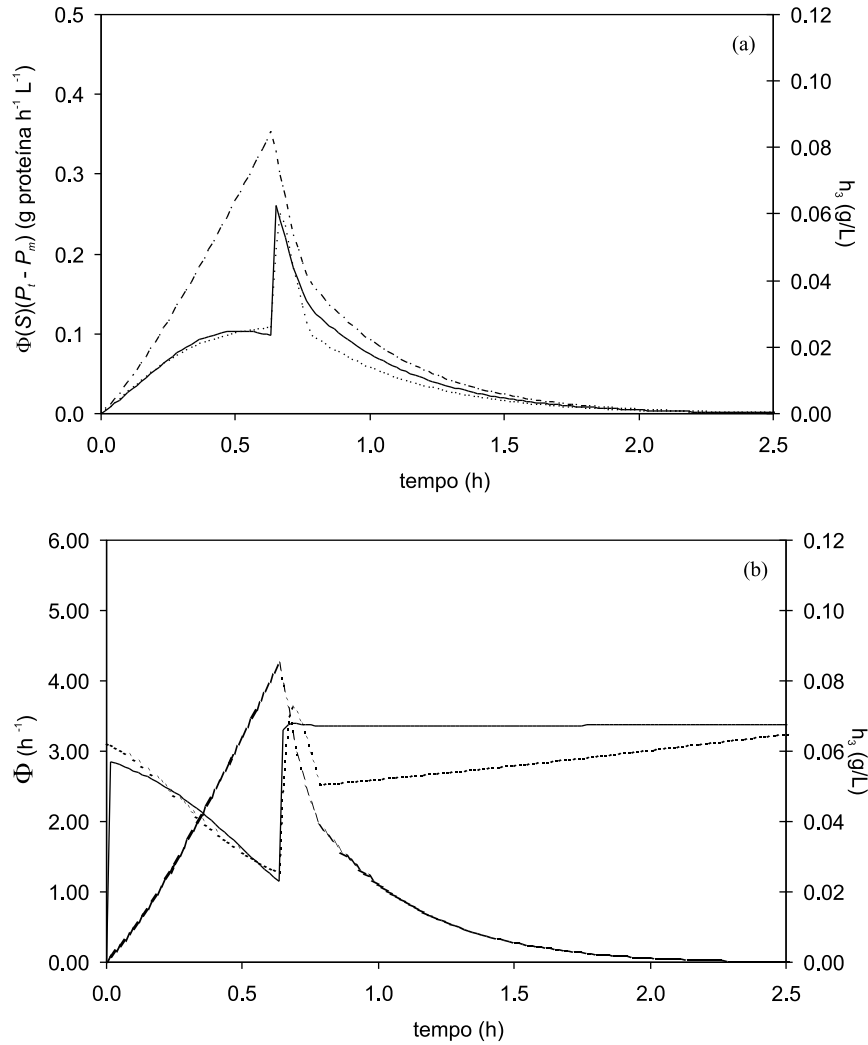


Figura 4.7: Cinéticas da secreção da proteína no meio: (a) taxa de secreção volumétrica; (b) taxa específica de secreção de proteína. As linhas a cheio representam as 'verdadeiras' funções e as linhas a tracejado representam as taxas do modelo. A linha ponto traço representa o h_3 .

Os resultados da identificação da taxa de secreção de proteína Φ foram surpreendentemente uma decepção. Este facto é inconsistente com os resultados excelentes que se obtiveram para a modelização das concentrações de P_m .

A análise da taxa de secreção volumétrica $v_{pm} = \Phi(S)(P_t - P_m)$ mostra um bom ajuste entre o modelo e os valores medidos (Figura 4.7(a)). Isto pode indicar que a função $\Phi(S)$ não é identificável por causa do termo $h_3 = (P_t - P_m)$.

Existe secreção de proteína só se a força motriz $(P_t - P_m)$ for diferente de zero. Em todos os estudos efectuados, $(P_t - P_m) > 0$ só para valores de concentração de

glucose baixos. Isto acontece porque a produção de proteína é inibida para valores de concentração de glucose elevados (repare-se no termo exponencial na Equação (3.7) e no gráfico 4.6).

Quando $f_p = 0$, P_t mantém-se constante e P_m converge rapidamente para P_t porque a equação $dP_m/dt = \Phi(S)(P_t - P_m)$ impõe uma convergência exponencial de P_m para P_t , assumindo que $\Phi(S)$ é positivo. Consequentemente, sempre que S é elevado $h_3 = (P_t - P_m) = 0$ e assim a condição C4 não é obedecida significando que a função $\Phi(S)$ não pode ser identificada.

Mesmo assim, para $h_3 = (P_t - P_m) > 0$ as saídas da rede neuronal estão de acordo com a função verdadeira como mostra a Figura 4.7(b) e isto explica porque é que as taxas de consumo volumétricas foram correctamente identificadas.

Também é importante realçar que quando P_t é constante a grandeza $\Phi(S)$ determina apenas a velocidade de convergência de $(P_m - P_t)$ e assim a função objectivo expressa pela Equação (4.18) para a identificação de $\Phi(S)$ será relativamente insensível para valores de concentração de glucose elevados.

4.6 Caso de Estudo II: processo de produção de fermento de padeiro

4.6.1 Descrição do processo

Adoptou-se neste capítulo o modelo matemático proposto por Sonnleitner e Kappeli (1986) para o crescimento da levedura *Saccharomyces cerevisiae* para estudos de simulação. Este modelo foi previamente descrito no Capítulo 3, na Secção 3.2.1.

O modelo dinâmico do biorreactor semicontínuo é obtido a partir do balanço material das componentes intervenientes no esquema reaccional (3.9) - (3.11) considerando que o reactor é perfeitamente agitado e os coeficientes de rendimento são constantes.

Por razões de simplicidade, assume-se que tanto as concentrações de O_2 como de N são mantidas acima dum nível limiar limitante por controlo. Além disso e uma vez que a concentração do CO_2 não interfere no metabolismo dos microrganismos (ver modelo cinético descrito no Capítulo 3, Secção 3.2.1) o conjunto das concentrações formado por X , S e E é suficiente para definir o estado do processo.

As equações de balanço material para estas três componentes podem ser escritas usando a seguinte representação matricial equivalente à Equação (4.3):

$$\frac{d}{dt} \begin{bmatrix} X \\ S \\ E \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ -k_1 & -k_2 & 0 \\ 0 & k_3 & -k_4 \end{bmatrix} \begin{bmatrix} \mu_s^o \\ \mu_s^r \\ \mu_e^o \end{bmatrix} X - D \begin{bmatrix} X \\ S \\ E \end{bmatrix} + \begin{bmatrix} 0 \\ DS_o \\ 0 \end{bmatrix} \quad (4.26)$$

Uma vez que o biorreactor é operado no modo semicontínuo, $D = F/V$, a equação de balanço global $dV/dt = F$ tem de ser integrada juntamente com a Equação (4.26).

Efectuaram-se seis corridas e simularam-se 16 h de operação com tempo de amostragem de 1 h. Os valores iniciais das concentrações de X e S foram gerados aleatoriamente a partir duma distribuição uniforme dentro da gama de 1–2 e 0–2 g/L, respectivamente. O valor inicial da concentração de etanol e do volume foram de 1 g/L e 1 L, respectivamente.

O perfil da taxa da alimentação F é representado por uma função em degrau gerada aleatoriamente a partir duma distribuição uniforme dentro da gama de 0–1 L/h e frequência de 2.7 h. O valor da concentração de glucose na corrente de alimentação foi de $S_o = 50$ g/L. Aos valores medidos de X , S e E foi adicionado erro Gaussiano com desvio padrão $\sigma = 0.25$.

4.6.2 Estrutura do modelo híbrido

Construíu-se um modelo híbrido considerando que a única parte do processo que é desconhecida no sentido mecanístico são as expressões cinéticas para o cálculo das taxas específicas de crescimento, μ_s^o , μ_s^r e μ_e^o . Assim, e de acordo com as definições apresentadas na Secção 4.2, a matriz das expressões cinéticas conhecidas é

$$\mathbf{H}(\mathbf{c}) = \text{diag} \left(\begin{bmatrix} X & X & X \end{bmatrix} \right)$$

e o vector das funções cinéticas desconhecidas é

$$\rho = \begin{bmatrix} \mu_s^o & \mu_s^r & \mu_e^o \end{bmatrix}^T.$$

O vector das funções $\rho(\mathbf{c})$ foi modelizado com uma rede MLP, como no caso anterior, com dimensão $\{2, 8, 3\}$ e factores de escala das saídas de 0.3 (estes têm de ser escolhidos ou heurísticamente por tentativa e erro ou pela análise dos valores medidos através da

Equação (4.14)) para todas as três taxas:

$$\rho = 0.3s(\mathbf{W}_2s(\mathbf{W}_1s(\mathbf{c}) + \mathbf{B}_1) + \mathbf{B}_2) \quad (4.27)$$

A rede tem duas entradas, S e E e o número total de parâmetros foi de $\dim(\mathbf{W}) = 51$, sendo \mathbf{W} a representação vectorial de \mathbf{W}_1 , \mathbf{B}_1 , \mathbf{W}_2 , \mathbf{B}_2 e a função de activação escolhida foi a função sigmóide expressa da seguinte forma $s(x) = 1/(1 + e^{-x})$. Assim, o modelo híbrido resume-se às Equações (4.26) e (4.27) com $n = 3$, $r = 3$ e $nw = 51$.

4.6.3 Resultados da identificação

Para a identificação de \mathbf{W} foram utilizados $P = 52$ dados que correspondem a três corridas do processo. Para evitar o sobreajustamento da rede foram usadas outras três corridas do processo para fazer validação cruzada (Pollard et al., 1992). Na técnica da validação cruzada, durante a optimização, o erro de validação é monitorizado juntamente com o erro de treino. A optimização é parada quando o erro de validação aumenta indicando que o modelo está a correlacionar-se com o ruído do conjunto de dados usado para treino. Adoptou-se a estratégia II para a identificação dos parâmetros do modelo híbrido porque o conjunto dos dados é esparso.

A optimização foi levada a cabo com um algoritmo SQP com os gradientes fornecidos pelas Equações das sensibilidades (4.20) - (4.22). O erro de validação (Figura 4.8) aumentou após 30 iterações ($CPU = 7476$ s) e a optimização foi parada. O erro quadrático médio final para os três conjuntos de dados usados para validação foi de 1.98×10^{-4} (com os valores das concentrações escalados pelo seu valor máximo 'medido' : 28.9 e 16.9 g/L para S e E , respectivamente).

Os resultados da modelização para o conjunto de dados de treino e para o conjunto de dados de validação são apresentados nas Figuras 4.9 e 4.10, respectivamente. Modelo e valores medidos mostram uma concordância excelente tanto no conjunto de dados de treino como no de validação. Neste caso a identificação foi bem sucedida para todas as seis corridas do processo.

A Figura 4.10 mostra os resultados obtidos com a rotina de integração de Runge-Kutta de 4^a/5^a ordem com tolerâncias absoluta e relativa de 1×10^{-5} e 1×10^{-3} , respectivamente, tanto para a integração das equações do modelo como para as equações das sensibilidades. O número total de ODEs neste caso é de 156, as quais têm de ser integradas entre todos os pontos medidos do conjunto de treino. Para conjun-

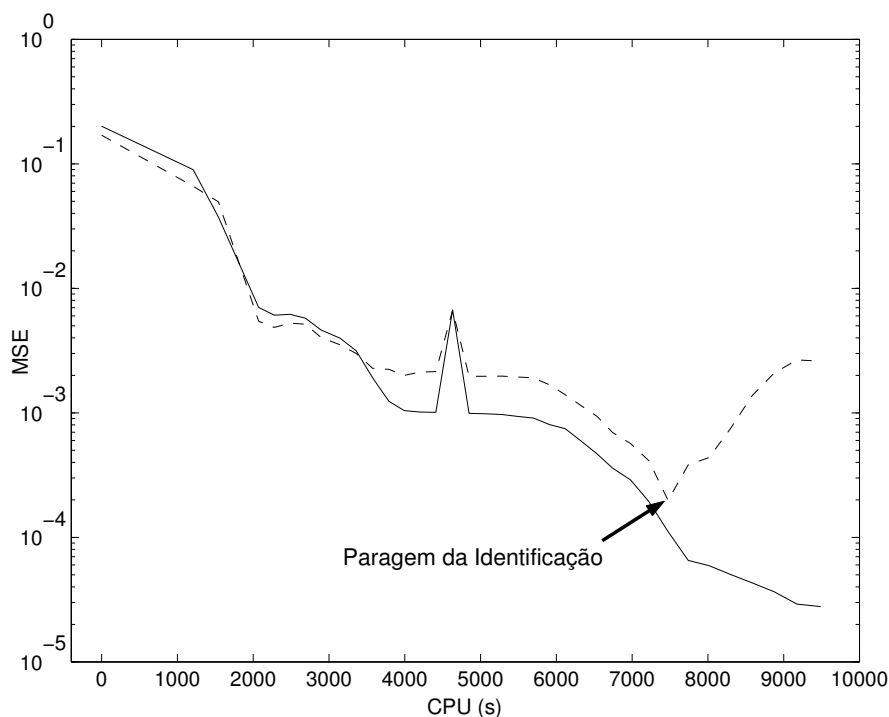


Figura 4.8: Erro quadrático médio (MSE) em função do tempo de computação (CPU, s), obtido num processador Pentium II: para a partição de treino (—, linha a cheio) e para a partição de validação (---, linha a tracejado)

tos de treino de grande dimensão e funções de rede grandes o tempo de computação pode ser um problema. Em princípio, a formulação discreta do modelo (4.3) - (4.4) é computacionalmente menos intensiva.

De facto, os modelos de redes neuronais dinâmicas são normalmente implementados na forma discreta (p.ex. Qin et al., 1992). Uma discretização exacta das equações de balanço material é descrita em Thompson e Kramer (1994) sob condições de simplificação, resultando num sistema linear dinâmico que poderá ser resolvido analiticamente. O modelo fornece previsões um passo à frente com taxas constantes descritas por redes neuronais embutidas no balanço material.

O modelo híbrido geral é não linear e por isso uma discretização exacta é impossível. Discretizações um passo à frente do modelo híbrido (4.3)- (4.4) foram derivadas empregando a técnica de discretização directa de Euler com períodos de amostragem de 0.1 h, 0.25 h e 1 h. A Figura 4.11 mostra os resultados obtidos para a concentração de glucose S para um único conjunto de dados de validação, isto é, só para uma corrida do processo.

Os resultados da modelização degradam-se à medida que o período de amostragem aumenta, o que de alguma forma era esperado. Contudo, foi observado que se se usar só

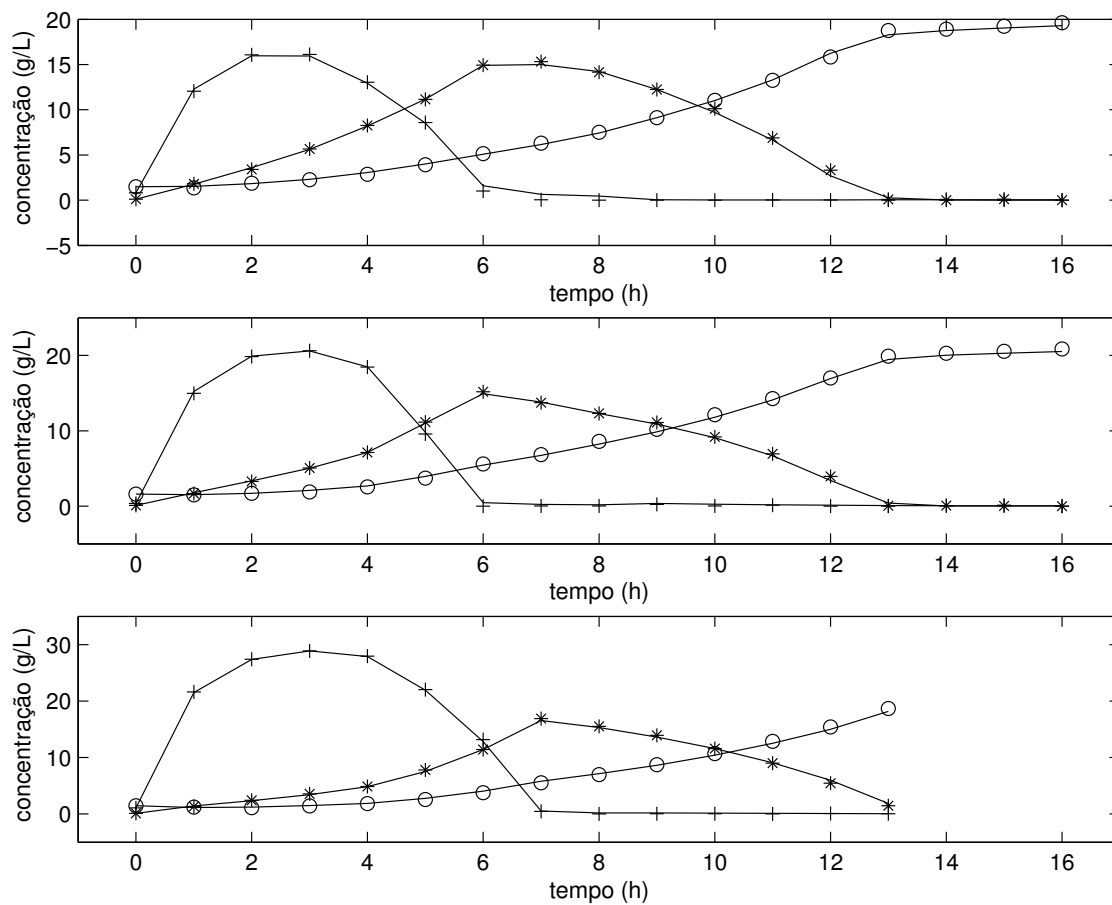


Figura 4.9: Resultados do modelo híbrido para o processo de produção do fermento de padeiro para os três conjuntos de dados usados para treino: (a) conjunto de treino 1; (b) conjunto de treino 2; (c) conjunto de treino 3. Os símbolos representam os valores 'medidos' e as linhas representam as saídas do modelo híbrido: +, concentração de glucose, *, concentração de etanol, o, concentração de biomassa.

um conjunto de dados para treino os resultados da modelização são exactos mesmo para o caso de $T = 1$ h. Neste caso o que parece acontecer é que para períodos de amostragem grandes o modelo híbrido perde a capacidade de descrever mais de um conjunto de dados simultaneamente. Isto pode ser explicado pelo facto de que na formulação discreta do balanço material as saídas da rede representam taxas médias estimadas sobre o período de amostragem. Obviamente, como nenhuma informação acerca da variação temporal das entradas - que depende de processos externos - é fornecida à rede, a relação já não é determinística.

O mesmo problema acontece quando se usa a estratégia I uma vez que as taxas estimadas pela Equação (4.14) ou pelo estimador OBE são valores médios sobre determinados intervalos de tempo.

O tempo de CPU obtido para $T = 1$ h foi de 88 s (ver Tabela 4.1) o que corresponde

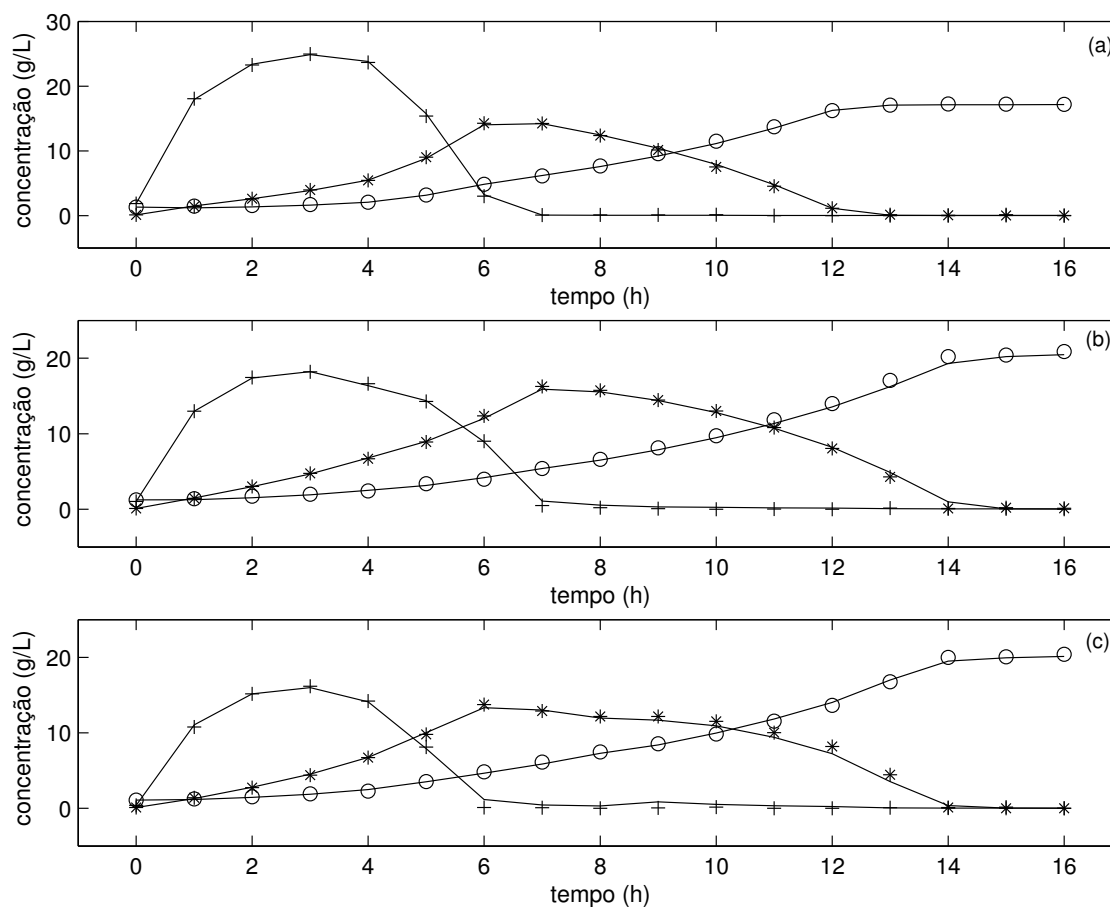


Figura 4.10: Resultados do modelo híbrido para o processo de produção do fermento de padeiro para os três conjuntos de dados usados para validação: (a) conjunto de validação 1; (b) conjunto de validação 2; (c) conjunto de validação 3. Os símbolos representam os valores 'medidos' e as linhas representam as saídas do modelo híbrido: +, concentração de glucose, *, concentração de etanol, o, concentração de biomassa.

a quase duas ordens de grandeza abaixo do obtido para o caso contínuo com integração de Runge-Kutta de 4^a/5^a ordem ($CPU = 7476$ s).

Tabela 4.1: Número de Iterações, Tempo de CPU, Erro de Validação e Erro de Treino em função dos métodos de discretização e contínuos utilizados

Métodos	Nº Iterações	CPU	MSE Treino	MSE Validação
RK	29	7476 s	1.11×10^{-4}	1.98×10^{-4}
$T = 0.1$ h	11	876 s	5.22×10^{-5}	2.52×10^{-4}
$T = 0.25$ h	6	332 s	2.32×10^{-4}	7.06×10^{-4}
$T = 1$ h	5	88 s	1.00×10^{-2}	2.37×10^{-2}

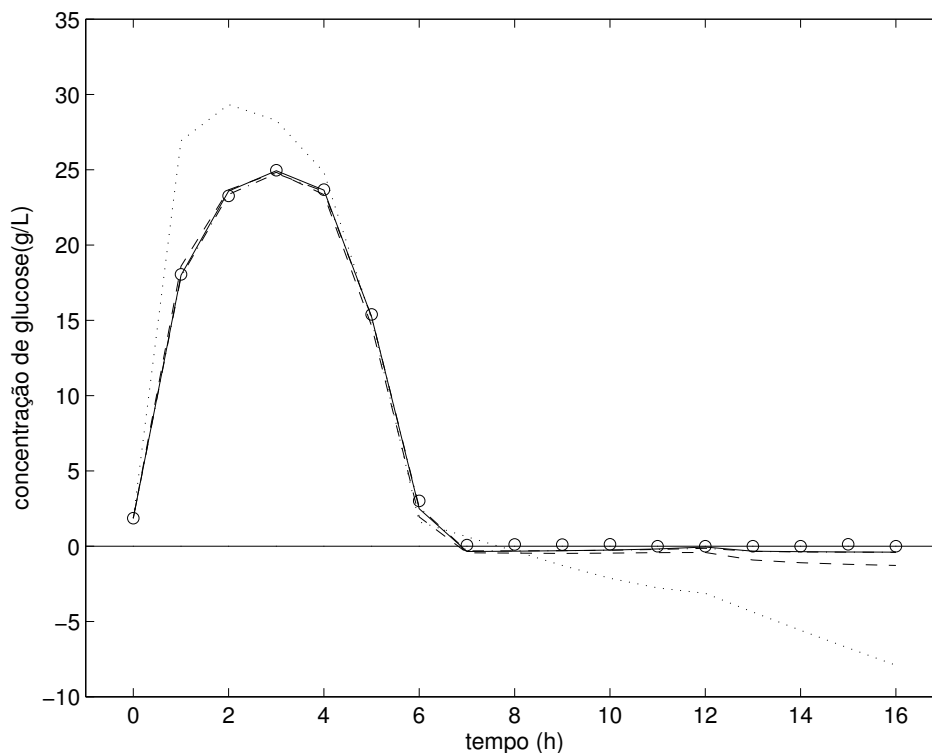


Figura 4.11: Resultados do modelo híbrido usando discretizações um passo à frente com períodos de amostragem de 0.1 (linha ponto traço), 0.25 (linha a tracejado) e 1 h (linha pontilhada), e usando a formulação contínua integrada com o algoritmo de Runge-Kutta de 4ª/5ª ordem (linha a cheio), símbolo o, valores medidos.

4.6.4 Resultados da estabilidade BIBO

Outro desenvolvimento importante é a caracterização do modelo híbrido dinâmico (4.3) - (4.4) em termos de estabilidade BIBO. No contexto deste caso de estudo verificou-se que os valores da modelização das concentrações de glucose eram por vezes negativos no fim do processo (Figura 4.12a). No entanto, valores de concentrações negativos não têm qualquer significado físico.

Este problema, típico nas técnicas de modelização não paramétricas, foi estudado na Secção 4.3 onde três condições de estabilidade BIBO foram apresentadas que garantem a positividade dos valores das concentrações assim como limites superiores. As condições C1 e C2 são dependentes do modo de operação do biorreactor e são válidas para este exemplo. No entanto, a condição C3 tem implicações na estrutura do modelo híbrido.

No caso em estudo, a condição C3 não é verificada com a definição prévia da matriz \mathbf{H} dada pela Equação (4.13). Contudo, a condição C3 é verificada se a matriz \mathbf{H} for

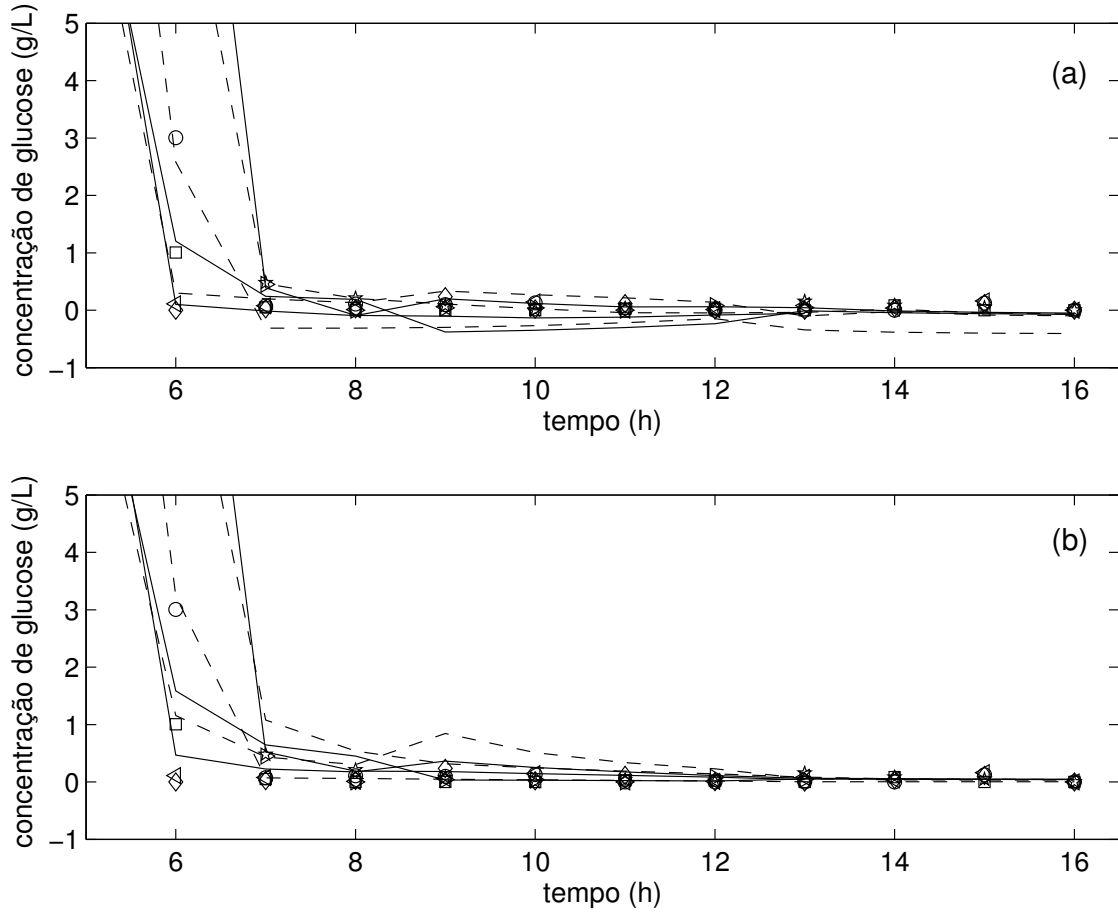


Figura 4.12: Teste das condições de estabilidade BIBO do sistema (4.3) - (4.4):(a) resultados da modelização da glucose para todos os conjuntos de dados com $\mathbf{H} = \text{diag}([X \ X \ X])$ e (b) resultados da modelização da glucose para todos os conjuntos de dados com $\mathbf{H} = \text{diag}([XS \ XS \ XE])$.

definida da seguinte forma de acordo com a mesma Equação (4.13):

$$\mathbf{H}(\mathbf{c}) = \text{diag}\left(\begin{bmatrix} XS & XS & XE \end{bmatrix}\right) \quad (4.28)$$

Esta definição também é consistente com as expressões cinéticas verdadeiras que entram no cálculo de μ_s^o , μ_s^r e μ_e^o definidas no Capítulo 3.

Desenvolveu-se um modelo híbrido novo com a matriz \mathbf{H} dada pela Equação (4.28). Assim sendo, modelizou-se o vector desconhecido ρ com uma rede similar ao caso anterior mas com factores de escala diferentes para as saídas além de que foram afinados heurísticamente:

$$\rho = \text{diag}\left(\begin{bmatrix} 9 & 0.9 & 2 \end{bmatrix}\right) s(\mathbf{W}_2 s(\mathbf{W}_1 s(\mathbf{c}) + \mathbf{B}_1) + \mathbf{B}_2) \quad (4.29)$$

em que $\dim(\mathbf{W}) = 51$ e $s(x) = 1/(1 + e^{-x})$. Identificou-se o modelo com os mesmos dados e métodos que no caso anterior. O erro de modelização obtido para o conjunto de validação foi igualmente baixo como no estudo anterior ($MSE = 1.24 \times 10^{-4}$ com concentrações escaladas pelos seus valores máximos). A diferença principal observada, como mostra a Figura 4.12, é que a concentração de glucose é agora estritamente positiva em conformidade com os resultados teóricos da Secção 4.3 e com a realidade física.

4.7 Conclusões

Melhoramentos na operação de bioprocessos implicam um trabalho cooperativo entre todos os factores que produzem conhecimento, i.é., cientistas, operadores de processo e técnicos. Daí que, quanto mais conhecimento existir acerca dos processos mais exactos podem ser os modelos dos processos e por sua vez mais eficientes podem ser as novas estratégias de operação baseadas em modelos.

De facto, a modelização híbrida pode constituir um veículo importante para a prossecução de tal filosofia de desenvolvimento porque representa uma maneira directa de ligar todos os factores que produzem conhecimento.

A maior fragilidade desta técnica é ainda a falta de uma base teórica sólida. Daí que, neste capítulo se tenha proposto uma definição para uma estrutura híbrida dinâmica geral válida para uma larga classe de problemas de modelização de biorreactores. A estrutura proposta permite a incorporação de diferentes fontes de conhecimento num sistema bem definido.

Discutiram-se duas estratégias para a identificação dos subsistemas embutidos. Provou-se que a melhor estratégia de identificação é a baseada na minimização directa dos erros nas concentrações no sentido dos mínimos quadrados empregando um algoritmo de SQP com os gradientes calculados pelas equações das sensibilidades. Concluiu-se que este método pode ser empregue com sucesso com dados experimentais esparsos e com ruído em oposição à estratégia I (que consiste na minimização dos erros nas taxas de reacção). O uso de métodos de integração robustos provou ser essencial para a identificação dos subsistemas desconhecidos embutidos na Equação (4.4). Observou-se também que a capacidade de descrever várias fermentações simultaneamente depende fortemente do grau de exactidão do método de integração empregue.

Discutiram-se também as condições de estabilidade BIBO para o modelo híbrido

dinâmico geral proposto. Propôs-se também uma nova configuração para o modelo híbrido que garante a positividade das concentrações em concordância com a realidade física.

Referências

- Bastin, G. e Dochain, D. (1990), *On-Line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam.
- Chen, L., Bernard, O., Bastin, G. e Angelov, P. (2000), 'Hybrid modelling of biotechnological processes using neural networks', *Control Engineering Practice*, **8**(7), 821–827.
- Dochain, D., Perrier, M. e Ydstie, B. E. (1992), 'Asymptotic observers for stirred tank reactors', *Chemical Engineering Science*, **47**(15-16), 4167–4177.
- Feyo de Azevedo, S., Dahm, B. e Oliveira, F. R. (1997), 'Hybrid modelling of biochemical processes: A comparison with the conventional approach', *Computers & Chemical Engineering*, **21**, S751–S756.
- Haykin, S. (1994), *Neural Networks: A comprehensive foundation*, Macmillan College Publishing Company, Inc.
- Haykin, S. (1999), *Neural Networks: A comprehensive foundation*, Prentice Hall, Inc., 2 Edição.
- Karama, A., Bernard, O., Gouze, J. L., Benhammou, A. e Dochain, D. (2001), 'Hybrid neural modelling of an anaerobic digester with respect to biological constraints', *Water Science and Technology*, **43**(7), 1–8.
- Leonard, J. e Kramer, M. A. (1990), 'Improvement of the backpropagation algorithm for training neural networks', *Computers & Chemical Engineering*, **14**(3), 337–341.
- Leonard, J. A., Kramer, M. A. e Ungar, L. H. (1992), 'A neural network architecture that computes its own reliability', *Computers & Chemical Engineering*, **16**(9), 819–835.
- Montague, G. e Morris, J. (1994), 'Neural-network contributions in biotechnology', *Trends in Biotechnology*, **12**(8), 312–324.

- Oliveira, R., Ferreira, E. C. e de Azevedo, S. F. (2002), 'Stability, dynamics of convergence and tuning of observer-based kinetics estimators', *Journal of Process Control*, **12**(2), 311–323.
- Park, S. e Ramirez, W. F. (1988), 'Optimal production of secreted protein in fed-batch reactors', *AIChE Journal*, **34**(9), 1550–1558.
- Peres, J., Oliveira, R. e de Azevedo, S. F. (2001), 'Knowledge based modular networks for process modelling and control', *Computers & Chemical Engineering*, **25**(4-6), 783–791.
- Pollard, J. F., Broussard, M. R., Garrison, D. B. e San, K. Y. (1992), 'Process identification using neural networks', *Computers & Chemical Engineering*, **16**(4), 253–270.
- Preusting, H., Noordover, J., Simutis, R. e Lubbert, A. (1996), 'The use of hybrid modelling for the optimization of the penicillin fermentation process', *Chimia*, **50**(9), 416–417.
- Psichogios, D. C. e Ungar, L. H. (1992), 'A hybrid neural network-1st principles approach to process modeling', *AIChE Journal*, **38**(10), 1499–1511.
- Qin, S. Z., Su, H. T. e Mcavoy, T. J. (1992), 'Comparison of 4 neural net learning-methods for dynamic system-identification', *IEEE Transactions on Neural Networks*, **3**(1), 122–130.
- Rumelhart, D. E., Hinton, G. E. e Williams, R. J. (1986), 'Learning internal representations by error propagation', em D. E. Rumelhart, J. L. McClelland e the PDP Research Group (Editores), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Cambridge, MA: MIT Press, volume 1: Foundations, Páginas 318 – 362.
- Schubert, J., Simutis, R., Dors, M., Havlik, I. e Lubbert, A. (1994a), 'Bioprocess optimization and control application of hybrid modeling', *Journal of Biotechnology*, **35**(1), 51–68.
- Schubert, J., Simutis, R., Dors, M., Havlik, I. e Lubbert, A. (1994b), 'Hybrid modeling of yeast production processescombination of a-priori knowledge on different levels of sophistication', *Chemical Engineering & Technology*, **17**(1), 10–20.
- Simutis, R., Oliveira, R., Manikowski, M., de Azevedo, S. F. e Lubbert, A. (1997), 'How to increase the performance of models for process optimization and control', *Journal of Biotechnology*, **59**(1-2), 73–89.

- Sonnleitner, B. e Kappeli, O. (1986), 'Growth of *Saccharomyces-cerevisiae* is controlled by its limited respiratory capacity formulation and verification of a hypothesis', *Biotechnology and Bioengineering*, **28**(6), 927–937.
- Thompson, M. L. e Kramer, M. A. (1994), 'Modeling chemical processes using prior knowledge and neural networks', *AIChE Journal*, **40**(8), 1328–1340.
- van Can, H. J. L., Braake, H. A. B. T., Hellinga, C., Luyben, K. C. A. M. e Heijnen, J. J. (1999), 'An efficient model development strategy for bioprocesses based on neural networks in macroscopic balances: Part ii', *Biotechnology and Bioengineering*, **62**(6), 666–680.
- van Can, H. J. L., teBraake, H. A. B., Dubbelman, S., Hellinga, C., Luyben, K. C. A. M. e Heijnen, J. J. (1998), 'Understanding and applying the extrapolation properties of serial gray-box models', *AIChE Journal*, **44**(5), 1071–1089.

Capítulo 5

Modelização de Cinéticas de Microrganismos com Mistura de Peritos

Conteúdo do Capítulo

No capítulo anterior propõe-se uma estrutura híbrida na qual as cinéticas da reacção são modelizadas em parte ou globalmente por redes neuronais. As redes que têm sido mais utilizadas nestes modelos híbridos são as redes Perceptrão de Camada Múltipla (MLP) e as redes de Funções de Base Radial (RBF). Neste capítulo propõe-se o uso de uma arquitectura complexa de redes modulares, chamada Mistura de Peritos (*Mixture of Experts* - ME), para modelizar cinéticas. A motivação para esta ideia reside no facto do metabolismo celular consistir em si mesmo numa rede modular complexa de reacções metabólicas. Este estudo foi elaborado com amostras simuladas de processos com vários níveis de complexidade e com dados experimentais dum processo de produção de fermento de padeiro. Como conclusões principais pode ser dito que a rede MLP e a rede mistura de peritos supera sistematicamente a rede RBF em termos do rácio exactidão do modelo/número de parâmetros. A rede modular ME treinada com o algoritmo da Máxima Verosimilhança é capaz de detectar as diferentes vias sem falhar e os peritos da rede desenvolveram individualmente sabedoria em descrever as vias metabólicas distintas. Em termos de exactidão, também a rede ME superou a rede MLP na sua habilidade para descrever a transição entre as vias metabólicas.

5.1 Introdução

As redes neuronais artificiais nomeadamente as redes Perceptrão de Camada Múltipla e as redes de Funções de Base Radial encontraram uma vasta aplicação em modelização de processos biológicos envolvendo catalizadores biológicos com sistemas celulares (Schubert et al., 1994; Montague e Morris, 1994; Feyerherm et al., 1997). Uma característica importante das células é o facto delas poderem processar substratos diferentes por vias metabólicas diferentes. Por exemplo, o crescimento diauxico na presença de duas fontes de carbono no meio, ou o crescimento aeróbio/anaeróbio dependendo da existência ou da ausência de oxigénio dissolvido no meio.

Concretamente, os exemplos seguintes contemplam processos biológicos com as características referidas atrás:

- (i) A levedura *S. cerevisiae* pode crescer por três vias metabólicas diferentes de modo a tirar partido de fontes energéticas e materiais básicas e é capaz de transitar entre um estado metabólico respirativo e um estado metabólico redutivo (Sonnleitner e Kappeli, 1986).
- (ii) Fornecendo excesso de glucose ou em condições de limitação de oxigénio a *Escherichia coli* produz através duma via metabólica alternativa acetato o que acima de determinadas concentrações inibe o crescimento (Reiling et al., 1985). No caso de haver limitação de glucose a *E. coli* é capaz de metabolizar acetato.
- (iii) Em casos mais complexos de culturas mistas vários mecanismos metabólicos diferentes podem ocorrer simultaneamente. Por exemplo, nos processos de tratamento de águas residuais por lamas activadas estão envolvidas várias populações de bactérias capazes de transitar entre estados metabólicos diferentes (Henze et al., 1999) tais como nitrificação/desnitrificação, aeróbio/anaeróbio, acumulação/libertação de fósforo.

Os sistemas biológicos acima exemplificados têm um crescimento cinético inerentemente não linear e descontínuo devido à transição entre os mecanismos metabólicos. Esta característica levanta algumas questões importantes no que diz respeito à modelização cinética com ANNs.

Com efeito, as redes MLP e RBF têm certas limitações para aproximar sistemas descontínuos de entrada/saída. Nomeadamente, as redes MLP têm tendência para exibir um comportamento irregular na vizinhança das descontinuidades (Haykin, 1994)

e as redes RBF estão mais vocacionadas para mapeamentos locais e não são muito apropriadas para a resolução de detalhes finos. Pelo contrário, as redes modulares podem ultrapassar os problemas descritos e têm potencial para modelizar as biocinéticas.

Neste capítulo faz-se um estudo comparativo da aplicação de redes neuronais para a modelização de biocinéticas, a saber: Redes de Percepção de Camada Múltipla, Redes de Funções de Base Radial e Redes de Mistura de Peritos. Na secção seguinte faz-se uma descrição sumária das redes de mistura de peritos.

5.2 Redes Mistura de Peritos

Os modelos de mistura (*mixture models*) da área de estatística (McLachlan e Basford, 1988; Titterington et al., 1985) inspiraram o desenvolvimento de alguns modelos na área das redes neuronais artificiais. Nomeadamente, uma classe de estruturas designadas por redes Mistura de Peritos (Jacobs et al., 1991), redes Mistura Hierárquica de Peritos (Jordan e Jacobs, 1994), assim como as Redes de Função de Base Radial Normalizadas (Moody e Darken, 1989).

Concretamente, num problema de regressão o objectivo é, dado um padrão $(\mathbf{x}_i, \mathbf{d}_i)$ onde \mathbf{x}_i pertence ao espaço das entradas S_{inp} e \mathbf{d}_i pertence ao espaço das saídas D_{out} , construir uma função $F_{me} : S_{inp} \rightarrow D_{out}$ que minimiza a superfície do erro de regressão, normalmente baseada no critério dos mínimos quadrados.

Assim sendo, a função de regressão da classe de estruturas baseada em modelos de mistura é definida da seguinte forma:

$$F_{me}(\mathbf{x}) = \sum_j P(j|\mathbf{x}) f_j(\mathbf{x}, \mathbf{w}_j) \quad (5.1)$$

em que \mathbf{x} é o vector de entrada, em que $f_j(\mathbf{x}, \mathbf{w}_j)$ é uma função de regressão local, o perito, definida pelos parâmetros \mathbf{w}_j . $P(j|\mathbf{x})$ é um peso não negativo de associação entre \mathbf{x} e o perito j que determina de forma efectiva o grau de contribuição de cada perito j para a saída final do modelo $F_{me}(\mathbf{x})$. Adicionalmente, é imposto que $P(j|\mathbf{x}) = 1$ o que naturalmente dá uma interpretação probabilística a este termo. Por sua vez, este termo é definido com sendo uma função paramétrica determinada por um conjunto de parâmetros \mathbf{a} .

Consequentemente, obtém-se a seguinte interpretação estatística do modelo. O par entrada/saída $(\mathbf{x}_i, \mathbf{d}_i)$ é gerado primeiro por amostragem aleatória de \mathbf{x}_i de acordo com

alguma função de densidade e depois por selecção aleatória dum modelo local de acordo com a função de probabilidade $P(j|\mathbf{x}_i)$. Assim sendo, para cada modelo local j , o seu vector de saída é uma variável aleatória cuja média é $f_j(\mathbf{x}_i, \mathbf{w}_j)$. Deste ponto de vista, $F_{me}(\mathbf{x})$ é interpretado como sendo o valor esperado da saída \mathbf{d} condicionada pela entrada \mathbf{x} .

Este tipo de modelos tem a vantagem de ser um compromisso entre modelos puramente locais por segmentos (*piecewise models*), como por exemplo as Árvores de Regressão e Classificação (*Classification and Regression Trees - CART*) (Breiman et al., 1984) e modelos globais como por exemplo MLPs. Os modelos puramente locais por segmentos dividem o espaço das entradas em sub-regiões de forma abrupta, onde cada sub-região é descrita exclusivamente pelo seu modelo. Efectivamente, a função de regressão por segmentos é composta por várias funções de regressão locais que cobrem colectivamente o espaço total das entradas. Além de particionar o espaço das entradas, o conjunto de parâmetros deste modelo é dividido em subconjuntos onde cada subconjunto só está activo numa determinada sub-região do espaço. Em contrapartida, nos modelos globais como as MLPs, existe apenas uma função de regressão que tem de ajustar os dados em toda a região do espaço das entradas onde não há uma divisão explícita deste espaço nem do conjunto de parâmetros.

De certo modo, os modelos puramente locais por segmentos como as CART são um caso particular dos modelos ME, em que $P(j|\mathbf{x})$ toma os valores 0 ou 1. Com efeito, os modelos ME também decompõem o problema de regressão na identificação dum conjunto de peritos, onde cada perito ajusta uma dada sub-região do espaço das entradas. No entanto nenhum dos peritos tem o exclusivo dessa região. A grande diferença entre estes dois tipos de modelos é que os modelos puramente locais por segmentos produzem aproximações que são descontínuas nas fronteiras das sub-regiões do espaço das entradas e os modelos ME são suaves em todo espaço das entradas devido à combinação linear definida em (5.1). Tipicamente, os métodos de aprendizagem empregues nos modelos puramente locais por segmentos são subóptimos devido à dificuldade de otimizar juntamente todos os seus parâmetros. No entanto, a aprendizagem nos modelos ME envolve naturalmente todos os parâmetros. Neste aspecto, estão mais próximos dos modelos globais como as MLPs em que a aprendizagem engloba todo o conjunto de parâmetros como por exemplo no método da retropropagação do erro ou noutros métodos baseados em gradientes descendentes.

Apesar de em problemas de regressão a função objectivo do algoritmo de treino ser definida geralmente pelo critério dos mínimos quadrados, Jacobs et al. (1991) e Jordan e Jacobs (1994) preferiram adoptar o critério baseado na máxima verosimilhança. Várias

razões levaram a esta escolha, nomeadamente, ter melhor desempenho e o treino ser mais rápido. Também a natureza da soluções influenciou esta escolha. De facto, concluíram que quando a função objectivo é definida pelo critério dos mínimos quadrados as soluções tendem a ser mais cooperativas enquanto que usando a formulação da máxima verosimilhança obtêm-se soluções mais competitivas. Isto é, enquanto que nos modelos cooperativos muitos peritos contribuem para uma dada saída, nos modelos competitivos essa contribuição é efectuada apenas por alguns peritos. A formulação deste problema baseada na máxima verosimilhança também é bastante atractiva pois permite aplicar o algoritmo da Esperança-Maximização (*Expectation-Maximization* - EM) (Dempster et al., 1977).

Jacobs et al. (1991) e Haykin (1994) também trataram o problema de estimação de parâmetros pela máxima verosimilhança no entanto empregaram o algoritmo do Gradiente Ascendente sendo estes gradientes calculados com o algoritmo de retropropagação. Todavia, Jordan e Xu (1995) mostraram que o algoritmo da Esperança-Maximização proporciona uma convergência linear que é mais rápida que a convergência obtida com o algoritmo do Gradiente Ascendente juntamente com a formulação da máxima verosimilhança.

Por sua vez, Rao et al. (1997) desenvolveram um algoritmo de treino para os modelos ME baseado no critério dos mínimos quadrados mas adoptando um método de optimização baseado no método do recozimento determinístico (*Deterministic Annealing* - DA) (Miller et al., 1996). Esta escolha foi baseada na assunção de que os métodos de gradientes não são os mais adequados para optimizar a superfície dos quadrados dos erros devido ao facto desta superfície ter numerosos óptimos locais.

Os modelos ME têm sido extensivamente aplicadas para reconhecimento de padrões, nomeadamente para análise de imagem (Dailey e Cottrell, 1999; Hinton et al., 1995; Melin et al., 2005), reconhecimento da fala (Peng et al., 1996; Waterhouse, 1993), simulação molecular (Barlow, 1995), em problemas de classificação (Jacobs et al., 1991; Hu et al., 1997), na área do controlo (Jacobs e Jordan, 1993; Jordan e Jacobs, 1994) e em problemas de regressão (Jordan e Jacobs, 1994; Weigend et al., 1995; Waterhouse, 1993; Rao et al., 1997). Que a autora tenha conhecimento este tipo de redes nunca foram aplicadas no contexto de modelização híbrida dinâmica de bioprocessos.

Como já foi dito atrás, as cinéticas das reacções nos microrganismos são regidas por uma rede bastante complexa de reacções metabólicas que podem ser vistas como sendo compostas por um conjunto de módulos interligados representando diferentes vias metabólicas: glicólise, ciclo dos ácidos tricarboxílicos (ciclo TCA) e muitas outras. Por

isso, a estrutura modular da rede ME parece ser bastante compatível com a estrutura interna do subsistema modelo cinético. Um segundo ponto relevante a favor das redes ME é que elas ajustam melhor sistemas de entrada/saída descontínuos (Haykin, 1994). Estas características importantes indicam que este tipo de redes pode ser vantajoso para modelizar as cinéticas das reacções.

No âmbito do estudo efectuado neste capítulo adoptou-se a estrutura rede de Mistura de Peritos definida em Jacobs et al. (1991). Esta estrutura é descrita em detalhe nas secções seguintes.

5.2.1 Arquitectura da mistura de peritos

O modelo ME inicialmente proposto e mais bem estudado é a rede mistura de peritos, também conhecida por modelo associativo de mistura de Gaussianas (Jacobs e Jordan, 1991; Haykin, 1994). A arquitectura da rede mistura de peritos, esquematizada na Figura 5.1, consiste num conjunto de K peritos e uma unidade de integração (também designada de sistema de ponderação). Basicamente, a tarefa de cada perito j é a de

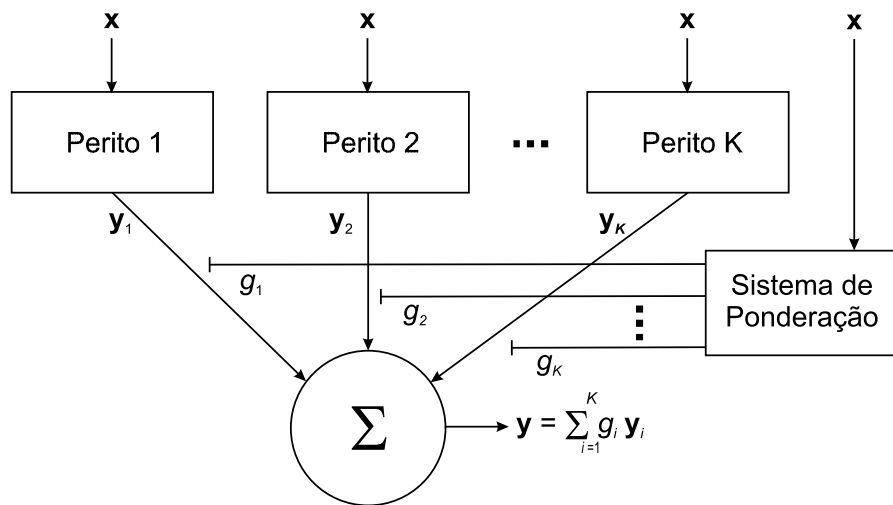


Figura 5.1: Diagrama de blocos da rede mistura de peritos: as saídas dos peritos são mediadas pelo sistema de ponderação

aproximar uma função $f_j : \mathbf{x} \rightarrow \mathbf{y}_j$ sobre uma região do espaço das entradas. A tarefa da unidade de integração é a de atribuir um perito a cada vector de entrada \mathbf{x} . A saída final do sistema \mathbf{y} é a combinação linear das saídas dos peritos e é expressa pela equação:

$$\mathbf{y} = \sum_{j=1}^K g_j(\mathbf{x}) \mathbf{y}_j(\mathbf{x}). \quad (5.2)$$

Conforme se refere na secção anterior esta estrutura tem fortes fundamentos estatísticos. Concretamente, a rede mistura de peritos definida pela expressão da Equação (5.2) pode ser vista em termos probabilísticos como sendo a probabilidade do padrão \mathbf{d} condicionada pela entrada \mathbf{x} representada pela expressão análoga seguinte:

$$P(\mathbf{d}|\mathbf{x}) = \sum_{j=1}^K P(j|\mathbf{x}) P(\mathbf{d}|\mathbf{x}, \mathbf{w}_j) \quad (5.3)$$

em que $g_j(\mathbf{x})$ é a $P(j|\mathbf{x})$ e obedece à seguinte restrição:

$$\sum_{j=1}^K g_j(\mathbf{x}) = 1, \quad g_j(\mathbf{x}) \geq 0. \quad (5.4)$$

e em que cada perito modeliza a função condicional de densidade de probabilidade $P(\mathbf{d}|\mathbf{x}, \mathbf{w}_j)$ do padrão alvo medido \mathbf{d} condicionado pela escolha do perito j e pelo valor de entrada \mathbf{x} . Daí que, as saídas dos peritos, $\mathbf{y}_j(\mathbf{x})$, sejam os valores médios das densidades $P(\mathbf{d}|\mathbf{x}, \mathbf{w}_j)$. Normalmente, para problemas de regressão assume-se que as densidades são Gaussianas (Jordan e Jacobs, 1994). Assumindo que as matrizes de covariância são da forma $\sigma_j^2 I$ resulta que:

$$P(\mathbf{d}|\mathbf{x}, \mathbf{w}_j) = \frac{1}{(2\pi)^{-n_d/2} \sigma_j^{n_d}} \exp\left(-\frac{1}{2\sigma_j^2} \|\mathbf{d} - \mathbf{y}_j\|^2\right) \quad (5.5)$$

em que \mathbf{d} tem dimensão n_d e o valor de entrada \mathbf{x} tem dimensão n_x .

Normalmente, os peritos são simples funções lineares para problemas de regressão não linear ou funções lineares com uma única saída não linear para problemas de classificação. Em alguns problemas de regressão não linear pode ser necessário usar peritos mais complexos por exemplo não lineares. Neste caso os peritos podem ser redes Perceptrão de Camada Múltipla e como se está perante um problema de regressão e não de classificação, estas redes devem ser definidas da seguinte forma (Bishop, 1995): funções de activação do tipo tangente hiperbólica nas camadas internas e função de activação linear na camada de saída. Este tipo de peritos foi adoptado neste trabalho e pode ser expresso da seguinte forma:

$$\mathbf{y}_j = \mathbf{W}_{2,j} \tanh(\mathbf{W}_{1,j} \mathbf{x} + \mathbf{b}_{1,j}) + \mathbf{b}_{2,j} \quad (5.6)$$

em que $\mathbf{W}_{1,j}$ e $\mathbf{W}_{2,j}$ são as matrizes dos pesos das ligações entre os nodos das camadas 1 e 2 e 2 e 3, respectivamente, e onde $\mathbf{b}_{1,j}$ e $\mathbf{b}_{2,j}$ são os vectores de parâmetros de desvio associados a cada camada da rede. Na análise que se segue os parâmetros

associados a cada perito j são representados vectorialmente da seguinte forma $\mathbf{w}_j = \{\mathbf{W}_{1,j}, \mathbf{b}_{1,j}, \mathbf{W}_{2,j}, \mathbf{b}_{2,j}\}$.

Na literatura encontram-se diferentes formas de definir a unidade de integração. A função *softmax* sugerida inicialmente por Jacobs et al. (1991) é uma função exponencial normalizada assegurando que as saídas do sistema de ponderação pertencem ao intervalo real $[0, 1]$. Esta definição estabelece uma divisão suave do espaço das entradas em hiperplanos (Ramamurti e Ghosh, 1999). Xu et al. (1995) definiram outro tipo de unidade de integração baseada em funções Gaussianas que divide o espaço das entradas em hiper-elipsóides e é designada na literatura por unidade de integração localizada ou Gaussiana. Este tipo de unidade de integração divide o espaço em regiões mais flexíveis (Ramamurti e Ghosh, 1999) e foi adoptada neste trabalho. Concretamente, pode ser expressa matematicamente da seguinte forma:

$$g_j(\mathbf{x}, \mathbf{a}_j) = \frac{\alpha_j P(\mathbf{x}, \mathbf{m}_j, \mathbf{\Sigma}_j)}{\sum_{i=1}^K \alpha_i P(\mathbf{x}, \mathbf{m}_i, \mathbf{\Sigma}_i)} \quad (5.7)$$

$$P(\mathbf{x}, \mathbf{m}_j, \mathbf{\Sigma}_j) = (2\pi)^{-n/2} |\mathbf{\Sigma}_j|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \mathbf{m}_j)^T \mathbf{\Sigma}_j^{-1} (\mathbf{x} - \mathbf{m}_j) \right\} \quad (5.8)$$

A Equação 5.8 é uma função de distribuição Gaussiana com centro \mathbf{m}_j e matriz de covariâncias $\mathbf{\Sigma}_j$ (de uma maneira geral só a diagonal da matriz das covariâncias é considerada). A expressão definida pela Equação (5.7) estabelece que as saídas da unidade de integração são normalizadas, obedecendo assim às restrições definidas pela Equação (5.4). Na Equação (5.7) a variável \mathbf{a}_j representa vectorialmente todos os parâmetros da unidade de integração, isto é, $\mathbf{a}_j = \{\alpha_j, \mathbf{m}_j, \mathbf{\Sigma}_j\}$.

Conceptualmente a arquitectura da rede mistura de peritos foi estendida de forma a incluir vários níveis hierárquicos sendo designada por rede Hierárquica de Mistura de Peritos (HME). A estrutura da rede HME é semelhante a uma árvore onde o espaço de entrada é subdividido em regiões que por sua vez são subdivididas em sub-regiões e assim sucessivamente. Esta estrutura é mais complexa e muitas vezes supera a versão não hierárquica da rede mistura de peritos (Haykin, 1999). Para o estudo presente a versão não hierárquica foi adoptada por razões de simplicidade visto que com ela já se obtiveram mapeamentos quase perfeitos e também porque Ramamurti e Ghosh (1999) concluíram que a versão não hierárquica com a unidade de integração localizada era semelhante à versão hierárquica.

5.2.2 O algoritmo da Esperança-Maximização

O algoritmo de aprendizagem baseado no método da máxima verosimilhança consiste na maximização da função de verosimilhança para o conjunto de dados medidos expressa por:

$$\begin{aligned}\mathcal{L}(\theta, \mathcal{X}) &= \prod_{t=1}^{n_p} \sum_{j=1}^K g_j(\mathbf{x}^t, \mathbf{a}_j) P(\mathbf{d}^t | \mathbf{x}^t, \mathbf{w}_j) \\ &= \prod_{t=1}^{n_p} \sum_{j=1}^K g_j(\mathbf{x}^t, \mathbf{a}_j) \frac{1}{(2\pi)^{-n_d/2} \sigma_j^{n_d}} \exp\left(-\frac{1}{2\sigma_j^2} \|\mathbf{d}^t - \mathbf{y}_j\|^2\right)\end{aligned}\quad (5.9)$$

em que $\theta = (\mathbf{w}_j, \mathbf{a}_j)$ e $\mathcal{X} = \{S_{inp}, D_{out}\}$.

Assim sendo, a maximização desta função produz estimativas de máxima verosimilhança para todos os parâmetros da rede de mistura de peritos, nomeadamente, de \mathbf{w}_j e \mathbf{a}_j . Normalmente, por questões práticas, a função que se optimiza é o logaritmo da função de verosimilhança dado por:

$$\ln \mathcal{L}(\theta, \mathcal{X}) = \sum_{t=1}^{n_p} \ln \sum_{j=1}^K g_j(\mathbf{x}^t, \mathbf{a}_j) \frac{1}{(2\pi)^{-n_d/2} \sigma_j^{n_d}} \exp\left(-\frac{1}{2\sigma_j^2} \|\mathbf{d}^t - \mathbf{y}_j\|^2\right) \quad (5.10)$$

A solução desta maximização pode ser obtida aplicando o algoritmo EM no entanto obriga a reformular o problema. Com efeito, na aplicação do algoritmo EM constata-se que a optimização da função $\mathcal{L}(\theta, \mathcal{X})$ seria simplificada se se conhecem um conjunto de variáveis omissas fictícias (*missing variable*). Assim sendo, definem-se estas variáveis designadas de indicadoras z_j^t do perito j para o padrão t assim:

$$z_j^t = \begin{cases} 1 & \text{se o padrão } t \text{ é gerado pelo perito } j \\ 0 & \text{se o padrão } t \text{ não é gerado pelo perito } j \end{cases}$$

Neste contexto, define-se o conjunto completo dos dados \mathcal{Y} como sendo o conjunto de dados $\mathcal{X} = \{S_{inp}, D_{out}\}$ mais o conjunto de variáveis omissas fictícias $\mathcal{Z} = \{z_j^t : j = 1, \dots, K, t = 1, \dots, n_p\}$. O algoritmo EM é utilizado para calcular os valores esperados para estas variáveis omissas. Na primeira parte do algoritmo, no passo-E, determina-se os valores esperados de z_j^t . Seguidamente, no passo-M, os parâmetros da rede, \mathbf{w}_j e \mathbf{a}_j são actualizados baseados nos valores esperados obtidos para z_j^t .

O algoritmo EM é um procedimento iterativo a dois passos onde cada iteração p pode ser resumida da seguinte forma (Jordan e Jacobs, 1994; Xu et al., 1995):

1. o passo-E:

O cálculo dos valores esperados para as variáveis indicadoras implica a definição da função de verosimilhança para o conjunto completo dos dados \mathcal{Y} que é dada pela expressão seguinte:

$$\mathcal{L}_c(\theta, \mathcal{Y}) = P(D_{out}, \mathcal{Z} | S_{inp}, \mathbf{w}_j, \mathbf{a}_j) = \prod_{t=1}^{n_p} \prod_{j=1}^K (g_j(\mathbf{x}^t, \mathbf{a}_j) P(\mathbf{d}^t | \mathbf{x}^t, \mathbf{w}_j))^{z_j^t} \quad (5.11)$$

em que $\theta = (\mathbf{w}_j, \mathbf{a}_j)$ representa os valores dos parâmetros da rede ME.

Na prática optimiza-se o logaritmo desta função dado pela expressão:

$$\begin{aligned} \ln \mathcal{L}_c(\theta, \mathcal{Y}) &= \sum_{t=1}^{n_p} \sum_{j=1}^K z_j^t \ln (g_j(\mathbf{x}^t, \mathbf{a}_j) P(\mathbf{d}^t | \mathbf{x}^t, \mathbf{w}_j)) \\ &= \sum_{t=1}^{n_p} \sum_{j=1}^K z_j^t (\ln g_j(\mathbf{x}^t, \mathbf{a}_j) + \ln P(\mathbf{d}^t | \mathbf{x}^t, \mathbf{w}_j)) \end{aligned} \quad (5.12)$$

Repare-se que com a introdução das variáveis omissas fictícias consegue-se passar o logaritmo para dentro do somatório (compare-se esta equação com a Equação (5.10)), o que simplifica substancialmente o problema de maximização.

Note-se também que a função $\mathcal{L}_c(\theta, \mathcal{Y})$ é uma variável aleatória dado que as variáveis \mathcal{Z} são na realidade desconhecidas. Por isso, a função que é optimizada é o valor esperado da verosimilhança completa Q dado o conjunto de dados medidos \mathcal{X} e o modelo actual definido pelos parâmetros $\theta^p = (\mathbf{w}_j^p, \mathbf{a}_j^p)$. Por conseguinte, define-se a função Q da seguinte forma:

$$Q(\theta, \theta^p) = E[\mathcal{L}_c(\theta, \mathcal{Y}) | \mathcal{X}] \quad (5.13)$$

donde

$$\ln Q(\theta, \theta^p) = \sum_{t=1}^{n_p} \sum_{j=1}^K h_j^{p,t} \left(\ln g_j(\mathbf{x}^t, \mathbf{a}_j^p) + \ln P(\mathbf{d}^t | \mathbf{x}^t, \mathbf{w}_j^p) \right) \quad (5.14)$$

onde se usa o facto de:

$$\begin{aligned}
 h_j^{p,t} &= E[z_j^t | \mathcal{X}] = P(z_j^t = 1 | \mathbf{d}^t, \mathbf{x}^t, \theta^p) \\
 &= \frac{P(\mathbf{d}^t | z_j^t = 1, \mathbf{x}^t, \theta^p) P(z_j^t = 1 | \mathbf{x}^t, \theta^p)}{P(\mathbf{d}^t | \mathbf{x}^t, \theta^p)} \\
 &= \frac{g_j(\mathbf{x}^t, \mathbf{a}_j^p) P(\mathbf{d}^t | \mathbf{x}^t, \mathbf{w}_j^p)}{\sum_{i=1}^K g_i(\mathbf{x}^t, \mathbf{a}_i^p) P(\mathbf{d}^t | \mathbf{x}^t, \mathbf{w}_i^p)} \quad (5.15)
 \end{aligned}$$

Concretamente, no passo-E calcula-se a matriz das probabilidades posteriores $\mathbf{h}^p = \{h_j^{p,t}\}$, assumindo densidades gaussianas, da seguinte forma:

$$h_j^{p,t} = \frac{g_j(\mathbf{x}^t, \mathbf{a}_j^p) \frac{1}{(2\pi)^{-n_d/2} \sigma_j^{n_d}} \exp\left(-\frac{1}{2\sigma_j^2} \|\mathbf{d}^t - \mathbf{y}_j\|^2\right)}{\sum_{i=1}^K g_i(\mathbf{x}^t, \mathbf{a}_i^p) \frac{1}{(2\pi)^{-n_d/2} \sigma_i^{n_d}} \exp\left(-\frac{1}{2\sigma_i^2} \|\mathbf{d}^t - \mathbf{y}_i\|^2\right)} \quad j = 1, \dots, K, t = 1, \dots, n_p \quad (5.16)$$

em que o subscrito j denota o índice do perito, o sobrescrito t refere-se a um padrão de treino e n_p representa o número de padrões medidos.

2. o passo-M, onde $K + 1$ problemas independentes de maximização são resolvidos. As primeiras K otimizações calculam os novos parâmetros \mathbf{w}_j de cada perito:

$$\mathbf{w}_j^{p+1} = \operatorname{argmax} \left(\sum_{t=1}^{n_p} h_j^{p,t} \ln P(\mathbf{d}^t | \mathbf{x}^t, \mathbf{w}_j^p) \right) \quad j = 1, \dots, K \quad (5.17)$$

Estas otimizações têm de ser resolvidas iterativamente usando métodos numéricos apropriados. Foi utilizado o método quasi-Newton com Gradientes Conjugados descrito em Moller (1993). Os gradientes foram calculados com retropropagação do erro nos peritos (Rumelhart et al., 1986).

A última otimização $K + 1$ calcula os novos parâmetros para a unidade de integração:

$$\mathbf{a}_j^{p+1} = \operatorname{argmax} \left(\sum_{t=1}^{n_p} h_j^{p,t} \ln g_j(\mathbf{x}^t, \mathbf{a}_j^p) \right) \quad j = 1, \dots, K \quad (5.18)$$

que no caso de se ter definido uma rede ME com a unidade de integração localizada (Equações (5.7) e (5.8)) a otimização tem uma solução analítica exacta obtida

num passo (Xu et al., 1995):

$$\alpha_j^{p+1} = \frac{1}{n_p} \sum_t h_j^{p,t} \quad (5.19)$$

$$\mathbf{m}_j^{p+1} = \frac{1}{\sum_t h_j^{p,t}} \sum_t h_j^{p,t} \mathbf{x}^t \quad (5.20)$$

$$\Sigma_j^{p+1} = \frac{1}{\sum_t h_j^{p,t}} \sum_t h_j^{p,t} (\mathbf{x}^t - \mathbf{m}_j^{p+1}) (\mathbf{x}^t - \mathbf{m}_j^{p+1}) \quad (5.21)$$

Os algoritmos descritos estão disponíveis na forma dum Toolbox do MatlabTM (Moerlan, 2000), o qual foi utilizado neste trabalho.

5.3 Caso de Estudo II: processo de produção de fermento de padeiro

5.3.1 Descrição do processo

O modelo matemático deste processo já foi descrito na Secção 3.2.1 do Capítulo 3. Relembramos alguns aspectos relevantes para o estudo efectuado nesta secção tendo presente o esquema reaccional dado pelas Equações (3.9)-(3.11).

As células do fermento de padeiro podem metabolizar glucose por duas vias metabólicas em determinadas condições aeróbias: via oxidativa e/ou via redutiva, sendo o etanol o produto final da via redutiva. As células do fermento de padeiro são também capazes de usar o etanol como fonte alternativa de substrato mas o etanol só pode ser metabolizado oxidativamente.

No estado metabólico oxidativo apenas as vias P1 e P3 (ver Equações (3.9)-(3.11)) estão envolvidas no processo de crescimento das células. Da análise do modelo resulta que o estado oxidativo ocorre para valores baixos de concentração de glucose nomeadamente abaixo de 0.042 g/L. O estado redutivo ocorre para concentrações acima de 0.042 g/L e corresponde às vias metabólicas P1 e P2. A transição entre estes estados metabólicos é uma transição rápida e portanto a via metabólica P2 e a P3 nunca ocorrem simultaneamente.

O objectivo neste estudo de caso é o de desenvolver um modelo neuronal que relacione a taxa específica de crescimento total (que é a soma das taxas específicas de crescimento

de cada via metabólica) com a composição do meio. Para simplificar a análise considerou-se que o oxigénio não era limitativo. Nestas condições a taxa específica de crescimento é apenas função das concentrações de glucose e etanol.

5.3.2 Resultados por simulação

Num primeiro estudo de simulação 6 corridas em semicontínuo foram simuladas usando o modelo descrito na Secção 3.2.1 do Capítulo 3, variando as condições de operação (taxa de alimentação, F , concentração de glucose na alimentação, S_o , e composição inicial do meio) segundo a Tabela 3.6 do capítulo referido, para as partidas B1 até B6.

Coleccionaram-se amostras da taxa específica de crescimento total em função da concentração de glucose S e da concentração do etanol E com intervalos de amostragem de 0.2 h. O número total de pontos obtido foi de $n_p = 606$. Estes dados foram usados para treinar e comparar as redes Mistura de Peritos, Perceptrão de Camada Múltipla e Funções de Base Radial.

A rede Mistura de Peritos foi configurada com $K = 2$ peritos. Os peritos são redes MLP de pequena dimensão definidas pela Equação 5.6 de dimensão $\{2, 2, 1\}$. As entradas da rede são S e E e a saída é a taxa específica de crescimento total, μ . O número total de parâmetros é de 24: 9 para cada perito e 6 para a unidade de integração (neste caso a unidade de integração só tem uma entrada, a concentração de glucose). A rede foi treinada com o algoritmo da Esperança-Maximização descrito na Secção 5.2.2.

Os resultados obtidos após 1000 iterações apresentam-se na Figura 5.2. Constata-se que a rede de mistura de peritos de pequena dimensão foi capaz de modelizar este sistema quase com erro negligenciável para todas as 6 corridas (o Erro Quadrático Médio (MSE) total obtido foi de 2.59×10^{-6}). Muito mais notável foi o facto do perito 1 ter-se especializado a descrever o estado metabólico oxidativo enquanto que o perito 2 se especializou a descrever o estado metabólico redutivo.

A Figura 5.3 mostra as saídas g_1 e g_2 da unidade de integração em função do número de pontos medidos juntamente com os valores da concentração de glucose medidos. As saídas g_1 e g_2 interceptam-se exactamente onde a transição entre os estados oxidativo e redutivo ocorre no processo verdadeiro que é para valores de concentração de glucose iguais a 0.042 g/L.

Uma rede MLP com 18 parâmetros foi treinada com os mesmos dados. A estrutura

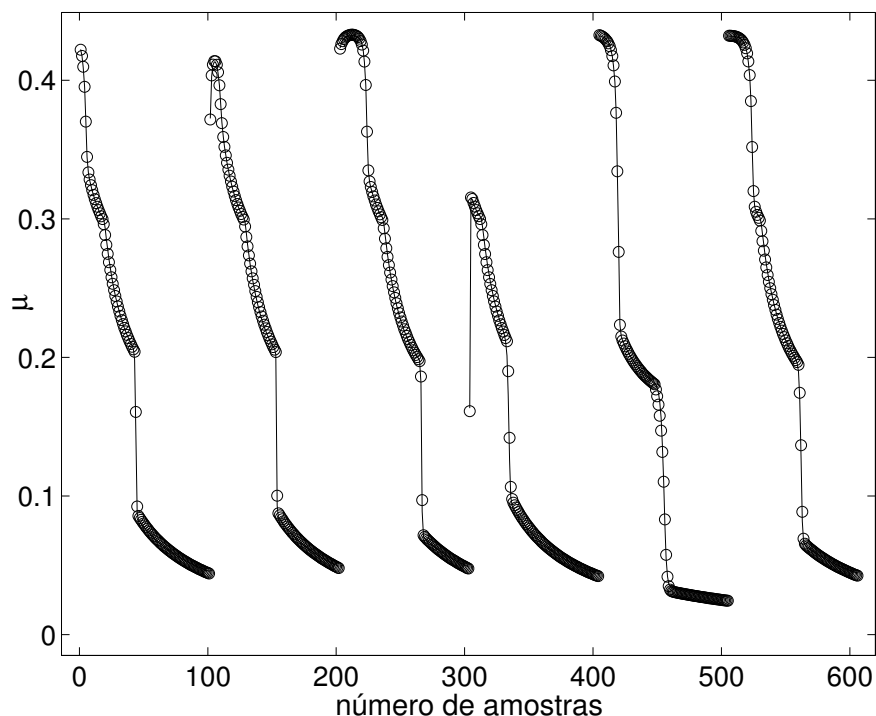


Figura 5.2: Resultados para 6 corridas simuladas: (a) taxa específica de crescimento estimada com uma rede ME com 2 peritos (18 parâmetros): valores medidos (o, pontos), valores estimados (-, linha)

desta rede é igual à estrutura dos peritos da rede mistura de peritos (Equação (5.6)). Esta rede foi treinada com o mesmo algoritmo empregue para resolver as primeiras K otimizações do passo M: o algoritmo quasi-Newton com o método do Gradiente Conjugado (CG) juntamente com retropropagação do erro para o cálculo analítico dos gradientes. Neste caso a função objectivo foi definida como um problema de mínimos quadrados. Após 2000 iterações o MSE estabilizou em 3.62×10^{-5} . Com efeito, este erro é superior ao erro obtido com a rede de peritos mas não é significativamente superior, diferindo só numa ordem de grandeza. Na prática obteve-se um mapeamento quase perfeito indicando não haver aparentemente vantagem em utilizar uma rede de mistura de peritos neste caso provavelmente por ser muito simples.

Os resultados obtidos com uma rede de Funções de Base Radial com as mesmas 2 entradas, S e E , uma única camada interna com 16 centros e uma camada de saída, treinada com os mesmos dados, foram bastante piores ($MSE = 2.3 \times 10^{-4}$). As funções de base radial escolhidas foram funções de densidade Gaussianas simétricas. O algoritmo de treino segue o esquema proposto por Moody e Darken (1989). Na primeira fase os centros e as larguras de cada função de base radial são determinados pelo método de agrupamento k-média e pela heurística P-vizinho-mais próximo, respectivamente. Na segunda fase, determinam-se os pesos das ligações entre a camada interna e a camada

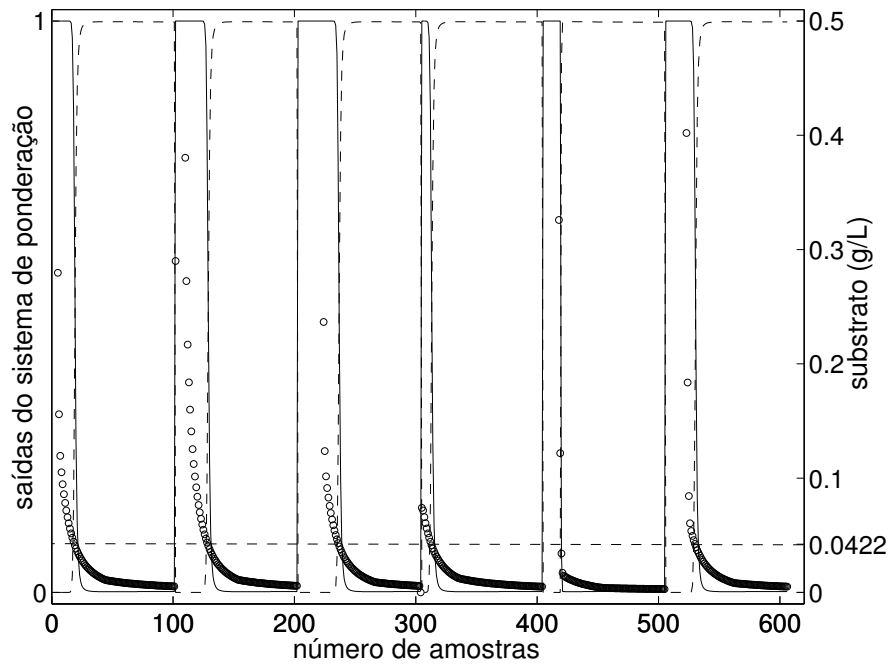


Figura 5.3: Resultados para 6 corridas simuladas: saídas da unidade de integração localizada: g_1 (...), g_2 (-, linha sólida) versus concentrações de S (o, pontos brancos). A verdadeira transição dá-se para substrato constante igual a 0.0422 (g/L)

de saída. Esta fase consiste simplesmente no cálculo duma matriz pseudo-inversa uma vez que a função do erro é quadrática nos pesos (Bishop, 1995).

5.3.3 Resultados experimentais

Utilizaram-se as experiências B1, B2, B3, B4 e B7 caracterizadas na Tabela 3.6 para fazer o estudo com dados experimentais. As amostras da biomassa X , da glucose S e do etanol E foram colecionadas em intervalos de tempo de aproximadamente 1.5 h. O valor da taxa específica de crescimento 'medida' foi calculada através da fórmula:

$$\mu_{exp} = \frac{1}{X} \frac{dX}{dt} + D \quad (5.22)$$

onde D denota a taxa de diluição ($D = F/V$, sendo V o volume do meio no fermentador). A derivada dX/dt foi calculada por splines cúbicos ajustando a biomassa ao longo do tempo seguido pela diferenciação analítica da função dos splines. A mesma rede de peritos do caso anterior foi agora treinada sobre os dados experimentais.

Os resultados obtidos são apresentados na Figura 5.4. Embora os erros de modelização obtidos, $MSE = 1.7 \times 10^{-3}$, sejam relevantes constata-se que as tendências

observadas nos dados experimentais foram capturadas pelo modelo de mistura de peritos. Os resultados do modelo podem até ser considerados razoáveis para as corridas B2, B3 e B4. No entanto, é importante realçar a existência de um certo grau de dispersão

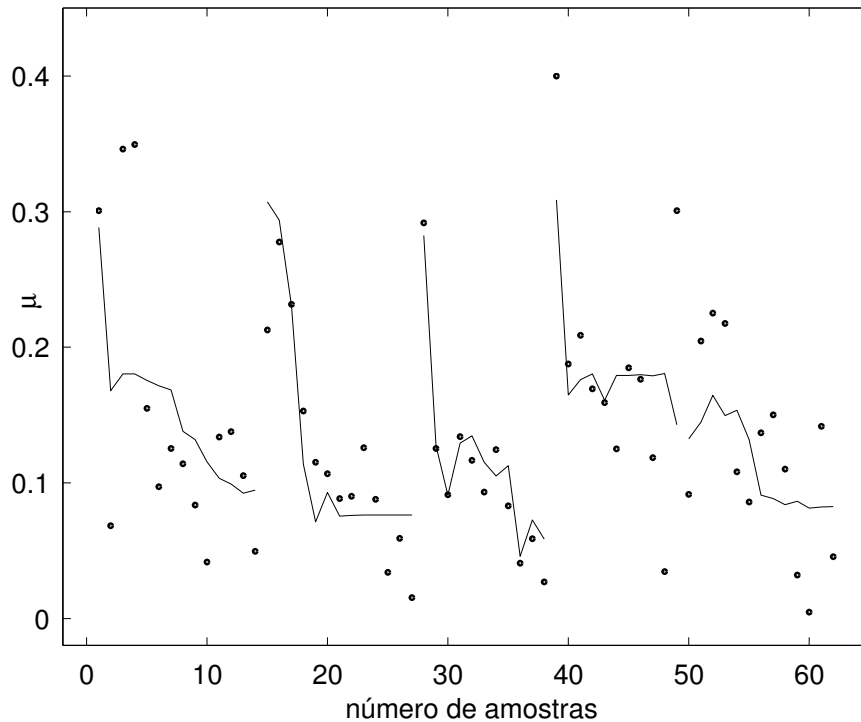


Figura 5.4: Resultados para 5 corridas com dados experimentais: (a) taxa específica de crescimento estimada com uma rede ME com 2 peritos (18 parâmetros): valores medidos (o, pontos), valores estimados (-, linha)

dos dados experimentais devido a erros de medida. Assim sendo, vários ajustes com splines de diferentes graus de suavização foram tentados mas foi impossível eliminar a dispersão sem perder informação intrínseca do processo. É também interessante observar que as simulações (Figura 5.2) e os dados experimentais (Figura 5.4) apresentam tendências similares no que diz respeito à variação da taxa específica de crescimento ao longo do tempo.

Contudo, o resultado mais importante é o facto da rede mistura de peritos conseguir de novo fazer a distinção entre o estado metabólico oxidativo e o estado metabólico redutivo como mostra a Figura 5.5. Esta figura representa as saídas da unidade de integração g_1 e g_2 sobre os pontos treinados juntamente com os valores da concentração de glucose medida (círculos). Pode-se constatar que $g_1 = 1$ ($g_2 = 0$) para baixas concentrações de glucose (abaixo de 0.1 g/L) enquanto que $g_2 = 1$ ($g_1 = 0$) para concentrações de glucose acima de 0.1 g/L indicando que a transição entre estados metabólicos ocorre para valores de $S = 0.1$ g/L. Este valor é maior que o valor teórico 0.042 g/L obtido

por simulação. O valor de transição é uma característica duma dada estirpe portanto não seria de esperar que se obtivesse o mesmo valor uma vez que o modelo apresentado na 3.2.1 do Capítulo 3 não foi ajustado às condições experimentais deste trabalho. Foi também observado que o valor de transição é muito sensível ao método usado para o cálculo da taxa específica de crescimento talvez devido à pouca qualidade dos dados experimentais. Claramente, neste exemplo os dados experimentais não permitem uma resolução suficiente à volta da transição, por isso o valor de 0.1 g/L deve ser considerado como um mero valor indicativo.

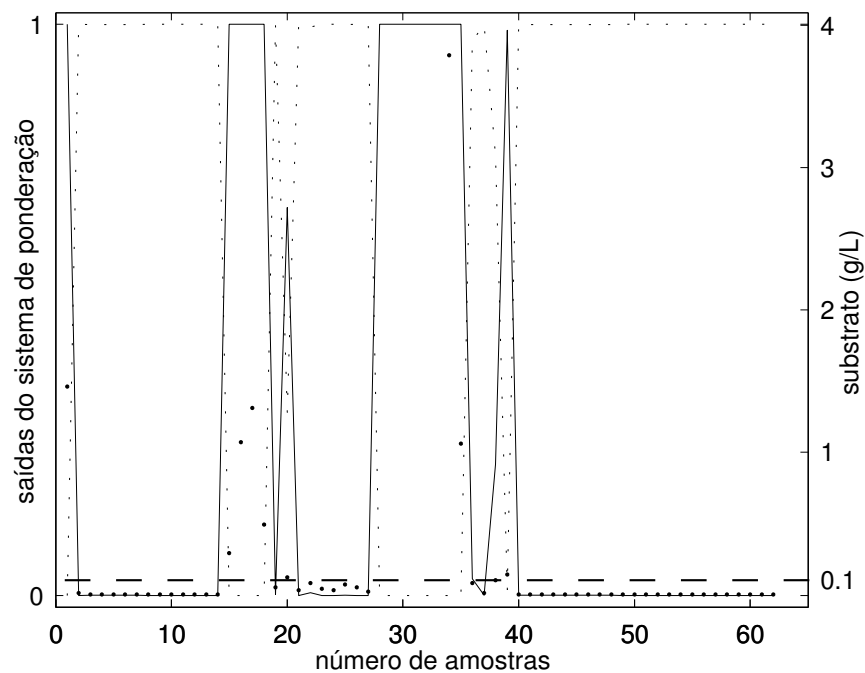


Figura 5.5: Resultados para 5 corridas com dados experimentais: saídas da unidade de integração localizada: g_1 (... , linha pontilhada), g_2 (-, linha sólida) versus concentrações de S (•, pontos pretos). A transição detectada dá-se para substrato na vizinhança de 0.1 (g/L)

Os resultados obtidos com um rede MLP e uma rede RBF de igual tamanho às da Secção 5.3.2 conduzem a um erro de $MSE = 1.9 \times 10^{-3}$ e $MSE = 1.67 \times 10^{-3}$, respectivamente. Estes erros são da mesma ordem de grandeza que o erro obtido com uma rede ME.

5.3.4 Erro na vizinhança da transição

É também relevante analisar o que acontece na vizinhança da transição. Para o efeito, gerou-se uma malha de 2601 pontos no plano das entradas formado por S e E usando

o modelo descrito na 3.2.1 do Capítulo 3. O objectivo é o de avaliar o desempenho de ambas as redes ME e MLP à volta da transição. A rede ME foi configurada com 2 peritos do tipo MLP cada um com a seguinte dimensão $\{2,2,1\}$. A saída da rede é a taxa específica de crescimento total e as entradas são $\mathbf{x} = \{S^t, E^t\}$. A unidade de integração é a unidade de integração localizada como no caso anterior (Equações (6.7) e (6.8)).

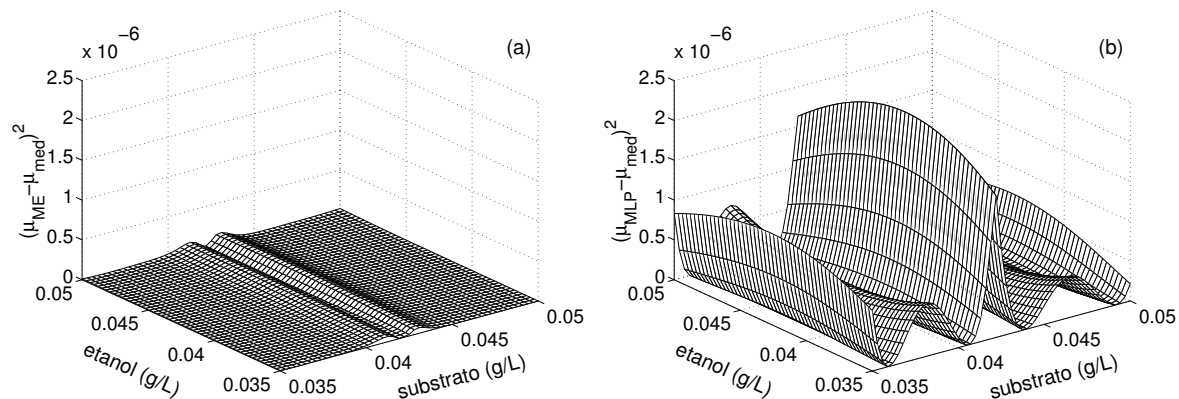


Figura 5.6: O quadrado do erro da estimação da taxa específica de crescimento com: (a) uma rede ME com 2 peritos MLP (18 parâmetros); (b) uma rede MLP com 17 parâmetros

Os resultados obtidos apresentam-se nas Figuras 5.6(a) e 5.6(b). As Figuras 5.6(a) e 5.6(b) mostram o erro de modelização para as redes ME e MLP respectivamente. Constatam-se duas diferenças evidentes:

- o quadrado do erro obtido para cada ponto é muito mais pequeno para a rede de Mistura de Peritos (ME) do que para a rede de Percepção de Camada Múltipla (MLP);
- o quadrado do erro obtido para cada ponto no caso da rede MLP é irregular.

De facto, este resultado é relevante no entanto não é totalmente inesperado pois é sabido que as redes MLP têm dificuldades a mapear sistemas descontínuos e exibem um comportamento oscilatório nas extremidades (Haykin, 1994). O uso da rede ME pode significar uma vantagem clara para a modelização de processos que correm perto da transição entre estados metabólicos. O caso da *S. cerevisiae* ou da *E. coli* são tais exemplos uma vez que estes microrganismos são aeróbios facultativos e a produção de etanol ou acetato está associada a baixos rendimentos de biomassa e de produto.

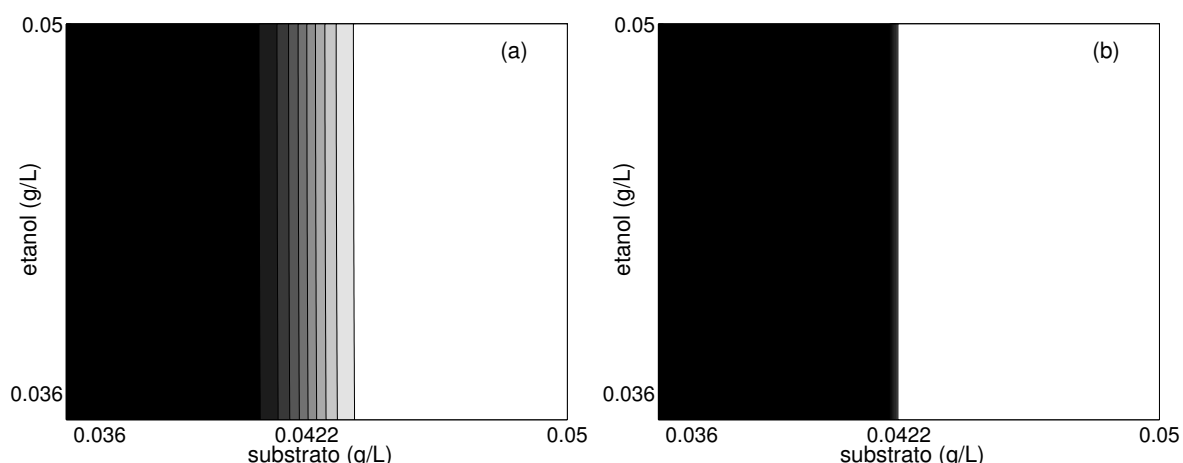


Figura 5.7: (a) Saídas da unidade de integração. (b) Verdadeira transição para substrato igual a 0.0422 (g/L)

A vantagem das redes ME em relação à rede MLP advém da capacidade de detectar a transição entre estados metabólicos e de atribuir a cada perito a descrição de cada estado metabólico individualmente. As Figuras 5.7(a) e 5.7(b) traduzem graficamente esta vantagem.

A Figura 5.7(b) mostra um desenho de curvas de nível para a verdadeira transição entre estados metabólicos do processo obtido directamente do modelo apresentado na Secção 3.2.1 do Capítulo 3 efectuando alguns cálculos. De facto, constata-se que a transição entre estados metabólicos é independente da concentração de etanol e ocorre exactamente para valores constantes de concentração de glucose $S = 0.042$ g/L. A cor preta representa o estado metabólico oxidativo e a cor branca representa o estado metabólico redutivo.

A Figura 5.7(a) é semelhante e representa as saídas da unidade de integração para o mesmo espaço de entradas $\mathbf{x} = \{S^t, E^t\}$. A cor preta representa a saída 1 da unidade de integração e a cor branca a saída 2 da unidade de integração. A transição ocorre precisamente para o mesmo valor da verdadeira transição, isto é, para $S = 0.042$ g/L. No entanto, a transição obtida pela rede de peritos é uma transição suave quando comparada com a verdadeira transição que é abrupta como mostra a Figura 5.7(b). Todavia, a suavidade da transição é uma característica dos sistemas biológicos e portanto não corresponde a uma desvantagem da unidade de integração localizada.

5.4 Caso de Estudo III: processo de remoção de fósforo de águas residuais por lamas activadas

5.4.1 Descrição do processo

Foi efectuado um estudo de simulação exaustivo baseado no modelo ASM2d (Henze et al., 1999). O modelo simplificado usado neste estudo de simulação já foi previamente descrito na Secção 3.3 do Capítulo 3. Neste modelo só são considerados os microrganismos designados por bactérias acumuladoras de fósforo e considera-se que o processo de operação consta de duas fases, a fase anaerobiose e a fase aerobiose.

Para este estudo, definiu-se que a primeira fase, a anaerobiose, tem uma duração total de 20 min por ciclo, sendo imediatamente seguida pela aerobiose que se definiu durar 10 min por ciclo. A transição entre as fases anaerobiose e aerobiose é imposta ligando ou desligando o arejamento.

Uma sequência de 13 partidas de um só ciclo foi simulada variando as concentrações iniciais do substrato fermentável, S_F , do acetato, S_A , do fosfato, S_{PO4} , do substrato lentamente biodegradável, X_S e da biomassa activa X_{PAO} . As condições iniciais foram perturbadas aleatoriamente em ciclos sucessivos, usando a distribuição uniforme dentro duma gama de $\pm 20\%$ à volta dos valores dados na Tabela 5.1, excepto para o caso do acetato que foi perturbado numa gama de $\pm 40\%$.

Os dados foram coleccionados com intervalos de amostragem de 0.005 h excepto para os 0.025 h iniciais de cada ciclo onde o intervalo de amostragem foi de 0.00125 h de forma a capturar mais informação sobre a influência do acetato no processo.

5.4.2 Modelos neuronais

Como já mencionado atrás a bactéria PAO é capaz de transitar entre estados aeróbios e anaeróbios. O objectivo neste estudo é o de analisar o desempenho duma rede ME para modelizar as cinéticas da bactéria PAO e em particular se é capaz de fazer a distinção entre os dois estados metabólicos. Como este exemplo é mais complexo que o anterior foi investigado se a rede ME supera a rede MLP não apenas numa vizinhança da transição mas sim em termos gerais.

Vários modelos neuronais foram treinados e comparados utilizando dados simulados. O objectivo da modelização é o de relacionar as concentrações S_{O2} , S_F , S_A , S_{PO4} , X_S

Tabela 5.1: Valores iniciais das variáveis de estado do modelo simplificado

Variável	Valor Médio	Descrição
S_{O_2}	0 g O_2 /m ³	concentração de oxigénio dissolvido
S_F	30 g COD /m ³	concentração de substrato fermentável
S_A	27 g COD /m ³	concentração de acetato
S_{NH_4}	1.26 g N /m ³	concentração de amónia
S_{PO_4}	0.9 g P /m ³	concentração de fosfato
S_I	30 g COD /m ³	concentração de fracção inerte
S_{ALK}	5 g HCO_3^- /m ³	alcalinidade em bicarbonato
X_S	125 g COD /m ³	concentração de substrato lentamente biodegradável
X_{PAO}	1800 g COD /m ³	concentração de biomassa activa
X_{PP}	450 g P /m ³	concentração de polifosfato armazenado
X_{PHA}	900 g COD /m ³	concentração de PHA armazenado
X_{GLY}	810 g COD /m ³	concentração de glicogénio armazenado

e X_{PAO} com as cinéticas específicas de consumo/produção correspondentes. A rede ME empregue está esquematizada na Figura 5.8. Esta rede foi configurada com 2 peritos porque o processo tem duas fases metabólicas. Neste caso também se adoptou por uma unidade de integração localizada previamente descrita na Secção 5.2.1.

Dada a complexidade e a não linearidade do processo os peritos MLP são de dimensão relativamente grande. Cada perito tem 5 entradas, $\mathbf{S} = \{S_{O_2}, S_F, S_A, S_{PO_4}, X_S\}$ e 6 saídas $\mathbf{q} = \{q_{SO_2}, q_{SF}, q_{SA}, q_{SPO_4}, q_{XS}, \mu\}$. O número de nodos internos foi variável. As funções de activação foram escolhidas como já descrito na Secção 5.2.1 para problemas de regressão não linear (Bishop, 1995): funções de activação tangente hiperbólicas para as camadas internas e funções de activação linear para camada de saída. No que diz respeito à unidade de integração localizada só a diagonal da matriz das covariâncias foi usada como sugerido por Ramamurti e Ghosh (1999).

As redes ME com peritos de tamanho variável (isto é, variando o número de nodos internos) foram sistematicamente comparadas com uma rede MLP de tamanho semelhante. Os parâmetros, tanto dos peritos MLP como da rede MLP, foram inicializados aleatoriamente a partir de uma Gaussiana com média zero e variância isotropica unitária onde a variância é escalada pelo número de nodos internos ou pelo número de saídas quando apropriado (Bishop, 1995). Os métodos de treino empregues foram o algoritmo EM no caso da rede ME e o algoritmo quasi-Newton com CG e retropropagação do erro no caso da rede MLP. Os dados simulados foram divididos em duas partes, uma parte de treino com 1160 pontos e uma parte para validação com 348 pontos.

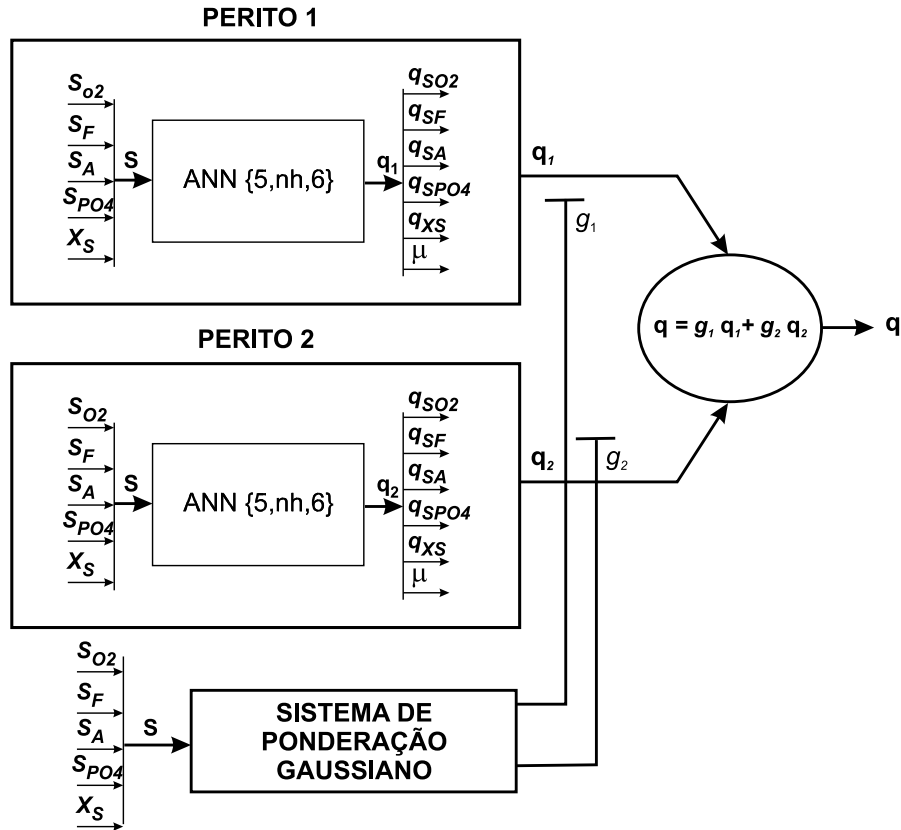


Figura 5.8: Estrutura da rede modular para modelizar as cinéticas de consumo/produção de S_{O_2} , S_F , S_A , S_{PO_4} , X_S e X_{PAO} com dois peritos MLP ($\mathbf{S} = \{S_{O_2}, S_F, S_A, S_{PO_4}, X_S\}$, $\mathbf{q} = \{q_{SO_2}, q_{SF}, q_{SA}, q_{SPO_4}, q_{XS}, \mu\}$, nh é o número de nodos da camada escondida)

5.4.3 Comparação entre as duas estruturas de rede

Foi feito um estudo exaustivo com o intuito de comparar o desempenho do ajuste feito por uma rede ME e uma rede MLP. As Figuras 5.9(a) e 5.9(b) apresentam o erro de modelização final para redes de dimensões diferentes, isto é, apresentam o erro em função do número total de parâmetros das redes consideradas. Cada ponto da Figura 5.9(a) foi obtido repetindo o procedimento de treino 30 vezes com parâmetros iniciais diferentes escolhidos aleatoriamente. Tanto no caso da rede ME como no caso da rede MLP, cada procedimento de treino consistiu em 1000 iterações seguido do cálculo do MSE para as partições de treino e validação.

A solução final corresponde ao valor mínimo do erro MSE obtido para a partição de validação nas 30 corridas. Os resultados finais apresentam-se nas Figuras 5.9(a) e 5.9(b) as quais contêm 3 curvas: os sinais + e os asteriscos referem-se à rede ME com ou sem incluir os parâmetros da unidade de integração, respectivamente, enquanto que os círculos referem-se à rede MLP.

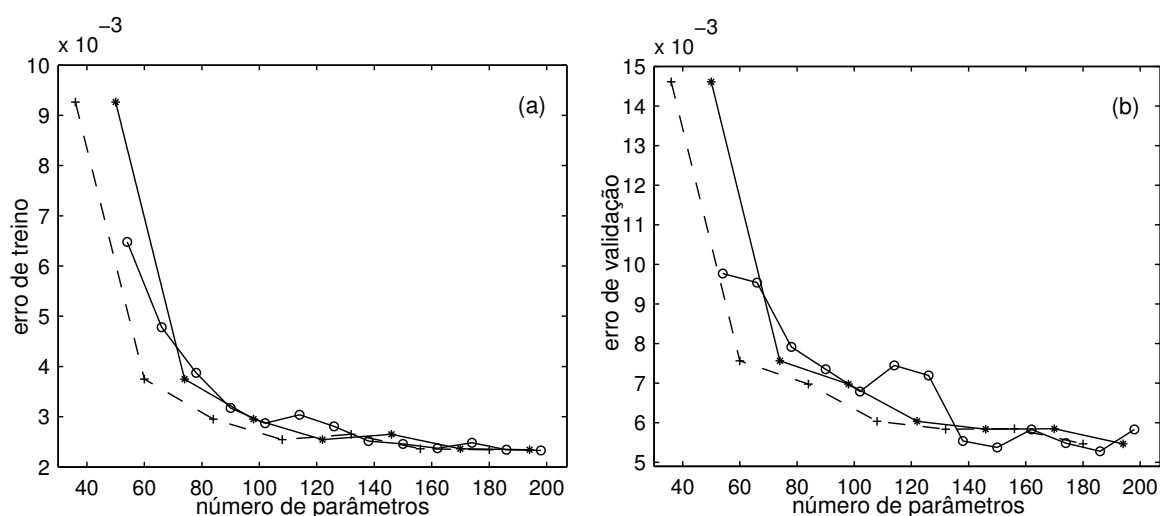


Figura 5.9: Erro quadrático médio obtido com uma rede MLP (o, pontos), com uma rede ME (*, asteriscos) e com uma rede ME sem incluir os parâmetros da unidade de integração (+, sinais mais): (a) partição de treino ;(b) partição de validação

Analisando a curva que não inclui os parâmetros da unidade de integração constata-se que independentemente do valor do número total de parâmetros os resultados obtidos pela rede ME são sempre mais exactos que os obtidos pela rede MLP. No entanto as duas curvas convergem para o erro de modelização mínimo à medida que o número de parâmetros aumenta. Provavelmente, quando o número total de parâmetros é grande a rede MLP apenas é menos exacta que a rede ME nas proximidades da transição o que no entanto no erro total tem um peso negligenciável.

Da comparação da curva que inclui os parâmetros da unidade de integração com a curva que não os inclui observa-se que os erros são superiores para a primeira curva quando o número total de parâmetros é pequeno. Este facto é facilmente explicável atendendo a que no início o número de parâmetros da unidade de integração corresponde a uma grande fracção do número total de parâmetros (14 parâmetros para a unidade de integração, 36 parâmetros para os peritos). Contudo, à medida que o número total de parâmetros aumenta o peso dos parâmetros da unidade de integração diminui e de novo as 3 curvas convergem para um mínimo do erro de modelização.

A Figura 5.10 apresenta o tempo de CPU gasto em função do número de parâmetros. É evidente que a rede ME tem uma desvantagem: o tempo de treino numa rede ME é dez vezes maior que o tempo de treino numa rede MLP. No entanto, tendo em consideração a evolução do poder computacional este facto não tem grande significado.

A característica mais importante e mais interessante das redes ME foi verificada em

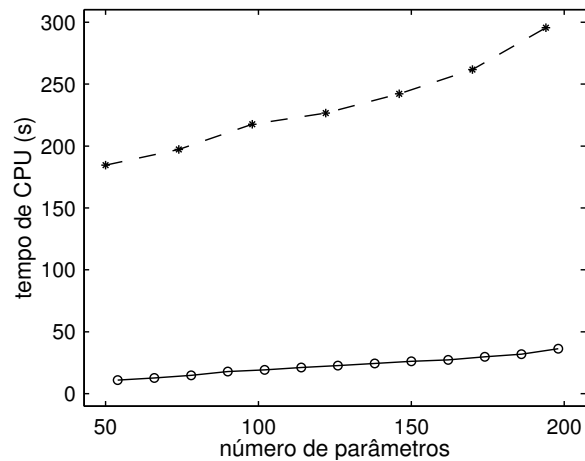


Figura 5.10: Tempo de CPU em segundos gasto no treino duma: rede MLP (o, pontos), rede ME (*, asteriscos)

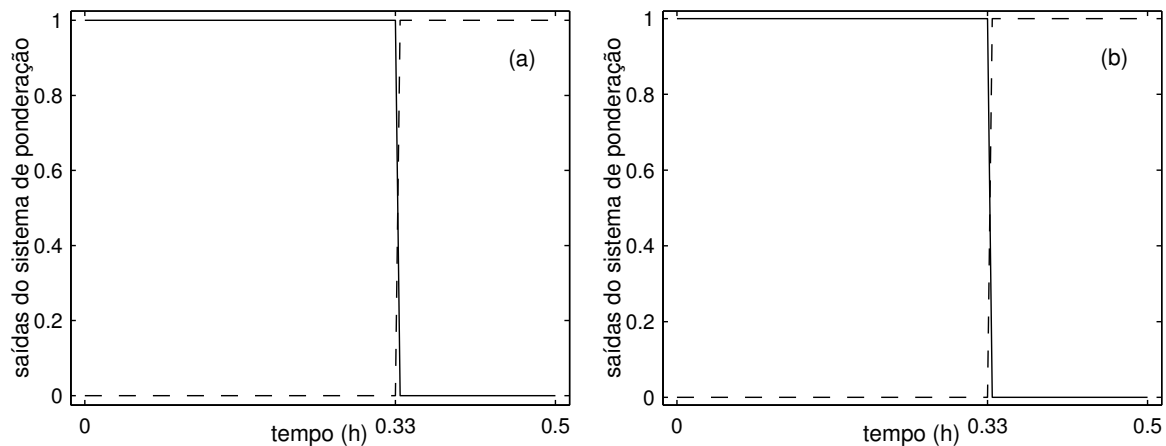


Figura 5.11: Saídas do sistema de ponderação: (a) partição de treino (-, linha a cheio);(b) partição de validação (-, linha a tracejado)

todos os estudos efectuados. A rede ME foi capaz de detectar a transição entre a anaerobiose e a aerobiose como mostram as Figuras 5.11(a) e 5.11(b) e, os peritos individualmente desenvolveram capacidades para modelizar as cinéticas de um ou do outro estado metabólico. A transição entre os peritos ocorre precisamente para o instante 0.33 h que corresponde à transição entre a anaerobiose e a aerobiose. Isto é observado para todas as corridas, tanto para as corridas de treino como para as corridas de validação.

Os resultados da modelização de ambas as redes ME e MLP para uma só corrida são apresentados na Figura 5.12. Modelos e medidas mostram um acordo notável no que diz respeito às cinéticas de S_{O_2} , S_A e S_{PO_4} . O mesmo não se pode dizer em relação às cinéticas de S_F , X_S e X_{PAO} . Este facto poderá ser explicado pelo exposto no parágrafo

seguinte.

Com efeito, o modelo utilizado neste capítulo inclui uma descrição do comportamento dinâmico de 3 grupos de compostos intracelulares, a saber, o PHA, o polifosfato e o glicogénio. No entanto, esta informação não foi incorporada nestes dois tipos de redes estáticas dada a sua natureza dinâmica. Provavelmente, este facto também explica porque é que os erros finais de modelização apresentados nas Figuras 5.9(a) e 5.9(b) estabilizam para valores elevados do erro mesmo quando se tem um grande número de parâmetros.

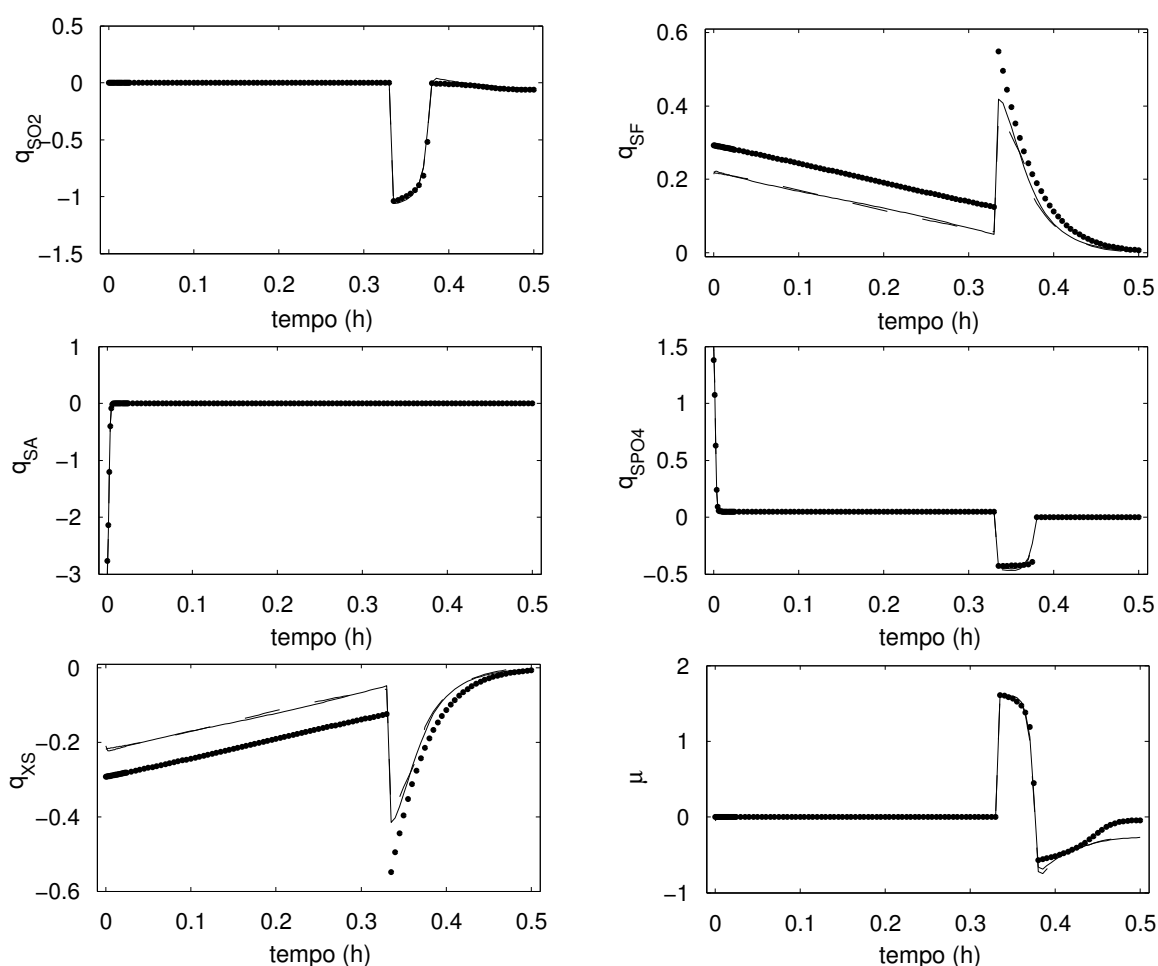


Figura 5.12: Resultados para uma partição de validação: valores medidos (o, pontos), rede MLP (-, linha sólida), rede ME (-, linha tracejada)

5.5 Conclusões

O principal objectivo do trabalho apresentado neste capítulo foi o de explorar a possibilidade de usar arquitecturas complexas de redes modulares para modelizar o modelo cinético em processos biológicos. Esta ideia é motivada pelo facto de o metabolismo das células em si consistir numa complexa rede modular de vias metabólicas.

Fez-se uma análise comparativa da rede mistura de peritos com as redes MLP e RBF habitualmente utilizadas na descrição do sistema célula no contexto da modelização híbrida. Esta análise foi suportada por dois casos de estudo com complexidades diferentes. Nomeadamente, estas redes foram utilizadas para modelizar a taxa específica de crescimento total num processo de produção de fermento padeiro, tanto a partir de dados simulados como de dados experimentais. No segundo caso de estudo, modelizaram-se as cinéticas de consumo/produção de oxigénio, de substrato fermentável, de acetato, de amónia, de substrato lentamente biodegradável e de biomassa activa num processo de remoção de fósforo de águas residuais por lamas activadas a partir de dados simulados.

As principais conclusões a que se chegaram neste estudo foram as seguintes:

- a rede Mistura de Peritos (ME) se treinada com o algoritmo Esperança-Maximização (EM) é capaz de detectar a transição entre estados metabólicos sem falhar;
- a rede Mistura de Peritos (ME) exhibe um desempenho comparável ao de uma rede Perceptrão de Camada Múltipla (MLP) em todos os testes elaborados;
- a rede Mistura de Peritos (ME) tem a vantagem adicional de os peritos empregues desenvolverem capacidades individuais para descrever os estados metabólicos individualmente;
- a rede Mistura de Peritos (ME) é capaz de descrever com mais exactidão as cinéticas na vizinhança das transições metabólicas.

Daí que, no contexto da modelização híbrida as redes modulares poderão representar um avanço na extracção de informação a partir de dados experimentais, produzindo modelos mais exactos e com melhor capacidade de extrapolação.

Referências

- Barlow, T. W. (1995), 'Feedforward neural networks for secondary structure prediction', *Journal of Molecular Graphics*, **13**(3), 175–183.
- Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, Oxford University Press.
- Breiman, L., Friedman, J. H., Olshen, R. A. e Stone, C. J. (1984), *Classification and Regression Trees*, Belmont, CA: Wadsworth.
- Dailey, M. N. e Cottrell, G. W. (1999), 'Organization of face and object recognition in modular neural network models', *Neural Networks*, **12**(7-8), 1053–1073.
- Dempster, A. P., Laird, N. M. e Rubin, D. B. (1977), 'Maximum likelihood from incomplete data via an algorithm', *Journal of The Royal Statistical Society Series B-Methodological*, **39**(1), 1–38.
- Feyo de Azevedo, S., Dahm, B. e Oliveira, F. R. (1997), 'Hybrid modelling of biochemical processes: A comparison with the conventional approach', *Computers & Chemical Engineering*, **21**, S751–S756.
- Haykin, S. (1994), *Neural Networks: A comprehensive foundation*, Macmillan College Publishing Company, Inc.
- Haykin, S. (1999), *Neural Networks: A comprehensive foundation*, Prentice Hall, Inc., 2 Edição.
- Henze, M., Gujer, W., Mino, T., Matsuo, T., Wentzel, M. C., Marais, G. V. R. e Van Loosdrecht, M. C. M. (1999), 'Activated sludge model no.2d, asm2d', *Water Science and Technology*, **39**(1), 165–182.
- Hinton, G., Revow, M. e Dayan, P. (1995), 'Recognizing handwritten digits using mixture of linear models', em G. Tesauro, D. Touretzky e T. Leen (Editores), *Advances in Neural Information Processing Systems*, The MIT Press, volume 7.

- Hu, Y. H., Palreddy, S. e Tompkins, W. J. (1997), 'A patient-adaptable ecg beat classifier using a mixture of experts approach', *IEEE Transactions on Biomedical Engineering*, **44**(9), 891–900.
- Jacobs, R. A. e Jordan, M. I. (1991), 'A competitive modular connectionist architecture', em J. M. R.P. Lippman e D. Touretzky (Editores), *Advances in Neural Information Processing Systems*, CA Morgan Kaufmann, San Mateo, volume 3, Páginas 767–773.
- Jacobs, R. A. e Jordan, M. I. (1993), 'Learning piecewise control strategies in a modular neural-network architecture', *IEEE Transactions on Systems Man and Cybernetics*, **23**(2), 337–345.
- Jacobs, R. A., Jordan, M. I., Nowlan, S. J. e Hinton, G. E. (1991), 'Adaptive mixtures of local experts', *Neural Computation*, **3**, 79–87.
- Jordan, M. I. e Jacobs, R. A. (1994), 'Hierarchical mixtures of experts and the em algorithm', *Neural Computation*, **6**(2), 181–214.
- Jordan, M. I. e Xu, L. (1995), 'Convergence results for the em approach to mixtures of experts architectures', *Neural Networks*, **8**(9), 1409–1431.
- McLachlan, G. H. e Basford, K. E. (1988), *Mixture Models: Inference and Application to Clustering*, New York: Marcel Dekker.
- Melin, P., Felix, C. e Castillo, O. (2005), 'Face recognition using modular neural networks and the fuzzy sugeno integral for response integration', *International Journal of Intelligent Systems*, **20**(2), 275–291.
- Miller, D., Rao, A. V., Rose, K. e Gersho, A. (1996), 'A global optimization technique for statistical classifier design', *IEEE Transactions on Signal Processing*, **44**(12), 3108–3122.
- Moerlan, P. (2000), *Mixture Models for Unsupervised and Supervised Learning*, Tese de Doutoramento, Computer Science Department, Swiss Federal Institute of Technology at Lausanne (EPFL).
- Moller, M. F. (1993), 'A scaled conjugate-gradient algorithm for fast supervised learning', *Neural networks*, **6**(4), 525–533.
- Montague, G. e Morris, J. (1994), 'Neural-network contributions in biotechnology', *Trends in Biotechnology*, **12**(8), 312–324.

- Moody, J. e Darken, C. J. (1989), 'Fast learning in networks of locally-tuned processing units', *Neural Computation*, **1**, 281 – 294.
- Peng, F. C., Jacobs, R. A. e Tanner, M. A. (1996), 'Bayesian inference in mixtures-of-experts and hierarchical mixtures-of-experts models with an application to speech recognition', *Journal of the American Statistical Association*, **91**(435), 953–960.
- Ramamurti, V. e Ghosh, J. (1999), 'Structurally adaptive modular networks for nonstationary environments', *IEEE Transactions on Neural Networks*, **10**(1), 152–160.
- Rao, A. V., Miller, D., Rose, K. e Gersho, A. (1997), 'Mixture of experts regression modeling by deterministic annealing', *IEEE Transactions on Signal Processing*, **45**(11), 2811–2820.
- Reiling, H. E., Laurila, H. e Fiechter, A. (1985), 'Mass-culture of escherichia-colimedium development for low and high-density cultivation of escherichia coli-b/r in minimal and complex media', *Journal of Biotechnology*, **2**(3-4), 191–206.
- Rumelhart, D. E., Hinton, G. E. e Williams, R. J. (1986), 'Learning internal representations by error propagation', em D. E. Rumelhart, J. L. McClelland e the PDP Research Group (Editores), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Cambridge, MA: MIT Press, volume 1: Foundations, Páginas 318 – 362.
- Schubert, J., Simutis, R., Dors, M., Havlik, I. e Lubbert, A. (1994), 'Hybrid modeling of yeast production processescombination of a-priori knowledge on different levels of sophistication', *Chemical Engineering & Technology*, **17**(1), 10–20.
- Sonnleitner, B. e Kappeli, O. (1986), 'Growth of *Saccharomyces-cerevisiae* is controlled by its limited respiratory capacity formulation and verification of a hypothesis', *Biotechnology and Bioengineering*, **28**(6), 927–937.
- Titterington, D. M., Smith, A. F. M. e Makov, U. E. (1985), *Analysis of Finite Mixture Distributions*, New York: Wiley.
- Waterhouse, S. R. (1993), *Speech recognition using hierarchical mixture of experts*, Tese de Mestrado, Cambridge University Engineering Department, Trumpington Street, Cambridge CB2 1PZ, UK.
- Weigend, A. S., Mangeas, M. e Srivastava, A. N. (1995), 'Nonlinear gated experts for time series: Discovering regimes and avoiding overfitting', *International Journal of Neural Systems*, **6**(4), 373–399.

- Xu, L., Jordan, M. I. e Hinton, G. E. (1995), 'An alternative model for mixture of experts', em G. Tesauro, D. S. Touretzky e T. K. Leen (Editores), *Advances in Neural Information Processing Systems*, MIT Press, volume 7, Páginas 633–640.

Capítulo 6

Modelização Híbrida Balanço Material/Mistura de Peritos

Conteúdo do Capítulo

Neste capítulo estuda-se uma estrutura híbrida semelhante à do Capítulo 4, mas onde as descrições das cinéticas é feita por redes de mistura de peritos. No modelo híbrido genérico, proposto no capítulo 4, as cinéticas são divididas numa parte mecanística e noutra do tipo caixa preta, a qual faz uma compensação da parte mecanística. Foram utilizadas essencialmente redes neuronais artificiais. Como se mostrou no Capítulo 5, o sistema 'cinética celular' possui algumas particularidades, nomeadamente uma estrutura intrínseca modular, a qual pode ser modelizada com vantagem usando redes mistura de peritos. Foi também demonstrado que, usando o algoritmo EM no treino destas redes, os peritos 'aprendem' a discriminar os diferentes estados metabólicos. Neste capítulo pretende estudar-se estruturas híbridas que combinam balanços materiais com redes de mistura de peritos. O treino destas redes juntamente com balanços materiais é agora mais complexo. As duas estratégias para identificação de parâmetros descritas no Capítulo 4 são adaptadas ao algoritmo EM e são comparadas. Estas metodologias foram usadas para modelizar um processo de produção de Polihidroxialcanoatos à escala laboratorial. Conclui-se que o modelo híbrido proposto foi capaz de descrever de modo exacto o comportamento dinâmico do processo e, além disso, o modelo foi capaz de se organizar em módulos que têm correspondência com as fases metabólicas da cultura.

6.1 Introdução

No Capítulo 4 propôs-se uma estrutura modelo híbrido genérica para biorreactores, dada pelas seguintes equações:

$$\frac{d\mathbf{c}}{dt} = \mathbf{v}(\mathbf{c}) - D\mathbf{c} + \mathbf{u} \quad (6.1)$$

sendo \mathbf{c} um vector de n concentrações, $\mathbf{v}(\mathbf{c})$ um vector de taxas volumétricas de reacção, D é a taxa de diluição e \mathbf{u} um vector de taxas volumétricas de entrada. O termo da reacção é dado por

$$\mathbf{v}(\mathbf{c}) = \mathbf{KH}(\mathbf{c})\rho(\mathbf{c}), \quad (6.2)$$

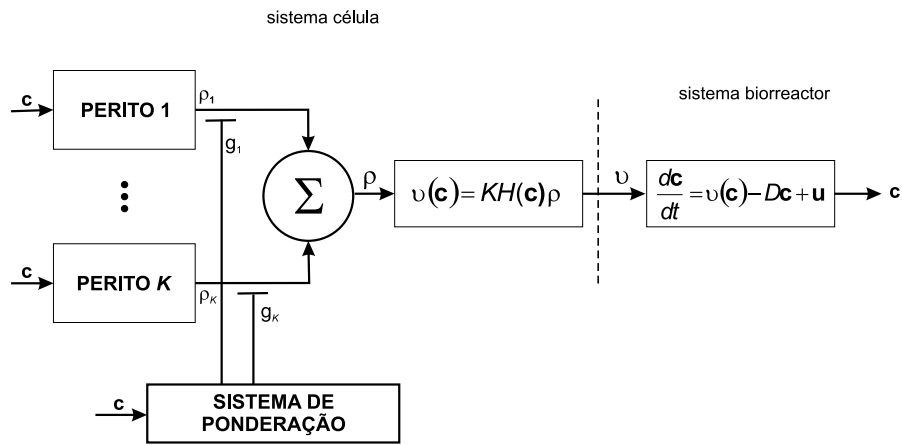
em que \mathbf{K} é uma matriz de coeficientes de rendimento de dimensão $n \times r$, $\mathbf{H}(\mathbf{c})$ é uma matriz de expressões cinéticas conhecidas, de dimensão $r \times r$, e $\rho(\mathbf{c})$ é um vector de r funções cinéticas desconhecidas.

No Capítulo 4 define-se $\rho(\mathbf{c})$ como uma rede neuronal do tipo MLP. No entanto, no Capítulo 5, mostrou-se que, devido à natureza dos sistemas biológicos em estudo, as redes de mistura de peritos podem ser mais adequadas para descrever as cinéticas em determinadas circunstâncias. Sendo assim, neste capítulo pretende desenvolver-se um modelo híbrido onde o termo $\rho(\mathbf{c})$ é dado por uma rede de mistura de peritos.

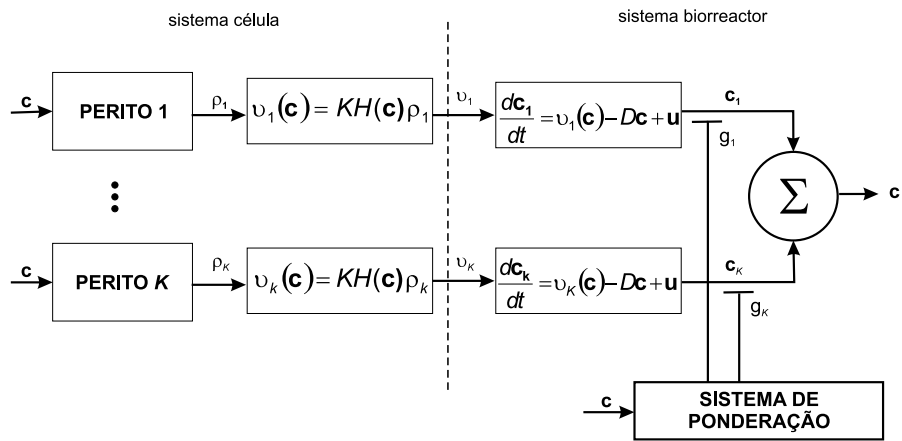
Genericamente, pretende-se desenvolver um modelo híbrido que integre os conceitos de rede ME e de dinâmica de sistemas, no caso na forma de equações de balanço material. Existem duas possibilidades para integrar estes dois conceitos numa estrutura híbrida (ver Figura 6.1(a,b)). As duas estruturas baseiam-se no princípio de competição de peritos e, além disso, seguem a estrutura genérica de modelo híbrido do Capítulo 4. É portanto introduzido um sistema de ponderação para regular a mistura de peritos. As duas estruturas diferenciam-se no ponto onde é efectuada a ponderação. No caso da Figura 6.1(a) a ponderação é efectuada nas cinéticas desconhecidas (ρ). No caso da Figura 6.1(b) a ponderação é efectuada na saída do modelo, i.e., nas concentrações.

Como se verá nas secções seguintes, a escolha de uma ou outra estrutura depende do método de identificação. No caso da estrutura da Figura 6.1(a) é possível usar-se a estratégia I descrita no Capítulo 4 enquanto que no caso da Figura 6.1(b) pode usar-se a estratégia II descrita no mesmo capítulo.

Deve notar-se que estas estruturas se reduzem à estrutura estudada no Capítulo 4 quando temos um único perito. Por isso, a estrutura do Capítulo 4 pode ser vista como



(a) Estrutura híbrida onde a ponderação é efectuada nas cinéticas



(b) Estrutura híbrida onde a ponderação é efectuada nas concentrações

Figura 6.1: Estrutura do modelo híbrido que combina redes ME com equações de balanço material.

um caso particular das estruturas propostas neste capítulo.

Este modelo vai ser validado experimentalmente utilizando o caso de estudo IV de produção de Polihidroxialcanoatos (PHA), nomeadamente ao tipo mais comum de PHAs que é o Poli- β -hidroxibutirato (PHB). A natureza biológica deste processo, descrita na Secção 3.4 do Capítulo 3, permite avaliar a aplicabilidade de modelos híbridos baseados em mistura de peritos, devido à ocorrência de estados metabólicos distintos.

6.2 Desenvolvimento dum modelo híbrido

6.2.1 Hipóteses simplificativas

Em culturas mistas existe um consórcio de bactérias que competem entre si pelos nutrientes existentes no meio. Pelo facto de existirem diversos tipos de microrganismos no meio, a modelização destes processos é caracterizada por modelos cinéticos complexos (ver os modelos de lamas activadas ASM1, ASM2, ASM2d e ASM3 compilados em Henze et al., 2000). O modelo desenvolvido neste trabalho assume as seguintes hipóteses simplificativas:

1. existência de apenas organismos heterotróficos;
2. dinâmica da cultura negligenciável. Assume-se que os organismos existentes no meio têm um comportamento metabólico médio que não se altera significativamente ao longo do tempo. Nestas circunstâncias a população de culturas mistas pode ser tratada como uma cultura homogénea.
3. acumulação aeróbia. O oxigénio dissolvido existe sempre em excesso e portanto o processo de acumulação intracelular mais importante é a acumulação aeróbia de COD na forma de PHB.
4. efeitos difusionais negligenciáveis. Assume-se agitação perfeita, com as células em suspensão, sendo a formação de flocos negligenciável. Logo, nem as resistências de transferência de massa externas nem as internas são consideradas no modelo.
5. meio definido quimicamente. Todos os nutrientes estão em excesso excepto o ácido acético e a amónia, que são os únicos substratos limitantes que podem ser usados para controlo.
6. modelo celular dividido em dois compartimentos. A massa das células é dividida em dois compartimentos: biomassa activa e PHB acumulado.

Destas hipóteses simplificativas destacam-se as mais importantes que são as da cultura homogénea com metabolismo médio invariável no tempo.

6.2.2 Equações de balanço material

Como foi já referido na Secção 3.4 do Capítulo 3, o reactor SBR é operado em ciclos de 'fome' e de 'fartura' onde no início de cada ciclo é adicionado novo meio de cultura

sendo o resto do ciclo operado essencialmente em modo descontínuo. Neste capítulo, desenvolve-se um modelo que descreve unicamente esta fase em descontínuo.

As equações de balanço material, que descrevem a fase descontínua do ciclo, tomam a seguinte forma na representação em espaço de estados:

$$\frac{d}{dt} \begin{bmatrix} X \\ HAC \\ NH4 \\ f_{PHB} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ -0.2 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} NH4X & 0 & 0 \\ 0 & HACX & 0 \\ -f_{PHB} & 0 & 1 \end{bmatrix} \begin{bmatrix} \rho(1) \\ \rho(2) \\ q_{PHB} \end{bmatrix} \quad (6.3)$$

com $\rho(1)NH4$ a taxa específica de crescimento, $\rho(2)HAC$ e q_{PHB} as taxas específicas de consumo de ácido acético e de consumo/formação de PHB , respectivamente. Neste modelo, assume-se que o rendimento amónia/biomassa ($Y_{N/X}$) é 0.2 N-mmol/C-mmol, definido a partir da fórmula empírica de biomassa proposta em Henze et al. (2000).

6.2.3 Estrutura do modelo híbrido

A estrutura do modelo híbrido foi derivada partindo do princípio que as equações de balanço material (Equação (6.3)) são conhecidas. A única parte desconhecida do processo do ponto de vista mecanístico são as cinéticas $\rho(1)$, $\rho(2)$ e q_{PHB} . Sendo assim, $\mathbf{u} = 0$, $D = 0$ e

$$\mathbf{c} = \begin{bmatrix} X & HAC & NH4 & f_{PHB} \end{bmatrix}^T$$

A matriz \mathbf{H} das expressões cinéticas conhecidas é

$$\mathbf{H} = \begin{bmatrix} NH4X & 0 & 0 \\ 0 & HACX & 0 \\ f_{PHB} & 0 & 1 \end{bmatrix}$$

e o vector das funções cinéticas desconhecidas é

$$\rho = \begin{bmatrix} \rho(1) & \rho(2) & q_{PHB} \end{bmatrix}^T.$$

A estrutura do modelo híbrido a adoptar depende da forma como os dados experimentais estão disponíveis, que por sua vez condicionam a estratégia de identificação de parâmetros. No Capítulo 4 descreveram-se duas estratégias para a identificação de pa-

râmetros. A estratégia I consiste na minimização dos erros nas cinéticas. Na estratégia II minimizam-se os erros nas concentrações. Na discussão que se segue, desenvolve-se a aplicação destas duas estratégias para as estruturas híbridas da Figura 6.1. O caso da estrutura da Figura 6.1(b) implica algumas alterações no algoritmo EM, como se descreve na Secção 6.3.2.

6.2.3.1 Estratégia I: minimização dos erros nas cinéticas

Pode adoptar-se a estratégia I quando os dados experimentais estão disponíveis em qualidade e quantidade suficientes, permitindo então estimar-se previamente as cinéticas desconhecidas. Esta estimativa é efectuada de forma indirecta a partir das medidas de concentrações ($\mathbf{c} = [X \text{ HAc} \text{ NH}_4 \text{ } f_{PHB}]^T$) usando a Equação (6.3). Para tal é necessário efectuar-se uma aproximação à derivada no tempo usando, por exemplo, o método das splines cúbicas (ver Secção 6.3.1). Com os valores 'experimentais' das cinéticas, é possível usar-se o algoritmo EM com ponderação nas cinéticas desconhecidas ρ , de acordo com a Figura 6.1(a).

Portanto, neste caso, os algoritmos de treino aplicados são os mesmos do Capítulo 5, isto é, o algoritmo EM 'standard'.

6.2.3.2 Estratégia II: minimização dos erros nas concentrações

A estratégia II usa-se quando os dados são esparsos, não sendo possível estimar-se as cinéticas de reacção. Opta-se então por uma minimização directa dos erros nas concentrações. Para que o pressuposto estatístico do algoritmo EM seja mantido, é necessário que a ponderação seja efectuada nas variáveis medidas, as quais possuem determinado modelo probabilístico associado (no caso concreto, o modelo gaussiano), ou seja, é necessário que a ponderação seja efectuada nas concentrações como se indica na estrutura da Figura 6.1(b). Isto implica, no entanto, que os cálculos associados a cada perito sejam alterados por forma que a diferença entre peritos seja 'mensurável' pelas concentrações de saída. Para tal, associa-se a cada perito um módulo de equações de balanço material como se indica na Figura 6.1(b). Assim, os cálculos associados a um perito j são da forma:

$$\frac{d\mathbf{c}_j}{dt} = \mathbf{v}_j(\mathbf{c}) - D\mathbf{c} + \mathbf{u} \quad (6.4)$$

$$\mathbf{v}_j(\mathbf{c}) = \mathbf{KH}(\mathbf{c})\rho_j(\mathbf{c}) \quad (6.5)$$

$$\rho_j = \mathbf{W}_{2,j} \tanh(\mathbf{W}_{1,j} \mathbf{c} + \mathbf{b}_{1,j}) + \mathbf{b}_{2,j} \quad (6.6)$$

em que $\mathbf{W}_{1,j}$ e $\mathbf{W}_{2,j}$ são as matrizes dos pesos das ligações entre os nodos das camadas 1 e 2 e 2 e 3, respectivamente, e $\mathbf{b}_{1,j}$ e $\mathbf{b}_{2,j}$ são os vectores de parâmetros de desvio associados a cada camada da rede. A rede ME foi configurada com dois peritos pois é sabido a priori que o processo tem duas fases metabólicas.

O sistema de ponderação gaussiano é definido da seguinte forma:

$$g_j(\mathbf{c}, \mathbf{a}_j) = \frac{\alpha_j P(\mathbf{c} | \mathbf{m}_j, \Sigma_j)}{\sum_{i=1}^K \alpha_i P(\mathbf{c} | \mathbf{m}_i, \Sigma_i)} \quad (6.7)$$

$$P(\mathbf{c} | \mathbf{m}_j, \Sigma_j) = (2\pi)^{-n/2} |\Sigma_j|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{c} - \mathbf{m}_j)^T \Sigma_j^{-1} (\mathbf{c} - \mathbf{m}_j) \right\} \quad (6.8)$$

em que $\mathbf{a}_j = \{\alpha_j, \mathbf{m}_j, \Sigma_j\}$.

De modo a incorporar estas alterações, o algoritmo EM foi redefinido do seguinte modo para cada iteração p :

1. o passo-E, onde a matriz das probabilidades posteriores $\mathbf{h}^p = \{h_j^{p,t}\}$ são calculadas pela expressão:

$$h_j^{p,t} = \frac{g_j(\mathbf{c}^t, \mathbf{a}_j^p) P(\mathbf{c}_{exp}^t | \mathbf{c}^t, \mathbf{w}_j^p)}{\sum_{i=1}^K g_i(\mathbf{c}^t, \mathbf{a}_i^p) P(\mathbf{c}_{exp}^t | \mathbf{c}^t, \mathbf{w}_i^p)} \quad j = 1, \dots, K, t = 1, \dots, n_p \quad (6.9)$$

em que o subscrito j denota o índice do perito, o sobrescrito t refere-se ao padrão de treino e n_p representa o número de pontos medidos. O termo $P(\mathbf{c}_{exp}^t | \mathbf{c}^t, \mathbf{w}_j^p)$ é a probabilidade condicional do padrão alvo medido \mathbf{c}_{exp} de dimensão n , dado o valor da entrada \mathbf{c} , de dimensão n , e dado o perito j . Assim,

$$P(\mathbf{c}_{exp}^t | \mathbf{c}^t, \mathbf{w}_j^p) = \frac{1}{(2\pi)^{n/2} |\Sigma_j|^{1/2}} \exp \left(-\frac{1}{2} (\mathbf{c}_{exp}^t - \mathbf{c}_j)^T \Sigma_j^{-1} (\mathbf{c}_{exp}^t - \mathbf{c}_j) \right) \quad (6.10)$$

em que \mathbf{c}_j é definido pelas Equações (6.4)-(6.6).

2. o passo-M, onde $K + 1$ problemas independentes de maximização são resolvidos. As primeiras K optimizações calculam os novos parâmetros \mathbf{w}_j de cada perito:

$$\mathbf{w}_j^{p+1} = \operatorname{argmax} \left(\sum_{t=1}^{n_p} h_j^{p,t} \ln P(\mathbf{c}_{exp}^t | \mathbf{c}^t, \mathbf{w}_j^p) \right) \quad j = 1, \dots, K \quad (6.11)$$

Estas optimizações foram resolvidas iterativamente usando o método quasi-Newton com Gradientes Conjugados onde os gradientes analíticos da função objectivo em ordem aos parâmetros foram calculados a partir das equações das sensibilidades (4.20)-(4.22) definidas no Capítulo 4. A última optimização $K + 1$ calcula os novos parâmetros para o sistema de ponderação gaussiano:

$$\alpha_j^{p+1} = \frac{1}{n_p} \sum_t h_j^{p,t} \quad (6.12)$$

$$\mathbf{m}_j^{p+1} = \frac{1}{\sum_t h_j^{p,t}} \sum_t h_j^{p,t} \mathbf{c}^t \quad (6.13)$$

$$\Sigma_j^{p+1} = \frac{1}{\sum_t h_j^{p,t}} \sum_t h_j^{p,t} (\mathbf{c}^t - \mathbf{m}_j^{p+1}) (\mathbf{c}^t - \mathbf{m}_j^{p+1}) \quad (6.14)$$

6.3 Resultados da identificação

Estudaram-se dois casos onde ocorre uma transição metabólica, a saber:

Caso 1: *NH4* como entrada no sistema de ponderação.

Neste caso a transição ocorre entre o estado de crescimento celular e o estado de não crescimento celular, i.e., quando $NH4 > 0$ ou $NH4 = 0$, respectivamente.

Caso 2: *HAc* como entrada no sistema de ponderação.

Neste caso a transição ocorre entre a fase de 'fome' ($HAc = 0$) na qual o metabolismo é redireccionado para o consumo das reservas intracelulares, e a fase da 'fartura' ($HAc > 0$), na qual as células crescem (se houver *NH4* no meio) e acumulam reservas intracelulares na forma de *PHB*.

Das 7 experiências seleccionadas para identificar os parâmetros, 5 delas foram utilizadas para o treino e 2 para a validação cruzada. A técnica da validação cruzada tem como objectivo evitar o sobre ajustamento do modelo aos dados como já referido no Capítulo 4. As 5 experiências usadas para o treino contêm 127 pontos e as 2 experiências utilizadas para validação contêm 51 pontos.

6.3.1 Estratégia I: minimização dos erros nas cinéticas

Primeiro foram obtidas estimativas das cinéticas ρ usando directamente a Equação (6.1) e empregando ajustes por splines cúbicas. As derivadas das concentrações em

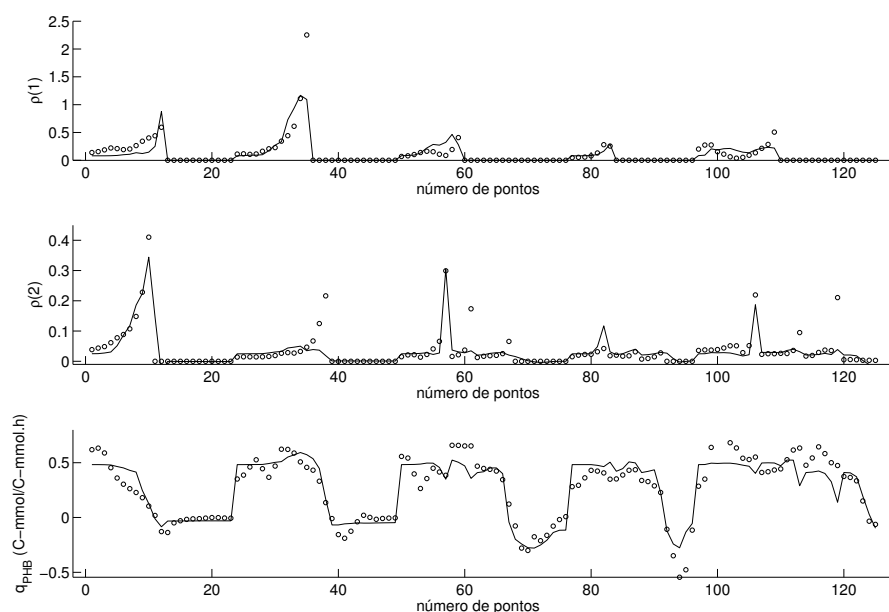


Figura 6.2: Resultados obtidos para $\rho(1)$, $\rho(2)$ e q_{PHB} para a partição de treino (caso1): valores experimentais (o, pontos), modelo ME (-, linha).

ordem ao tempo foram calculadas ajustando as concentrações por splines cúbicas seguido de diferenciação analítica das funções spline.

Cada rede MLP foi definida com 3 entradas (X , HAc e f_{PHB}), 3 nodos na camada interna e 3 saídas ($\rho(1)$, $\rho(2)$ e q_{PHB}). As entradas e as saídas de cada rede MLP são as mesmas da rede ME. A função de activação escolhida foi a função de activação tangente hiperbólica para a camada interna e a função linear para a camada de saída. O número total de parâmetros foi de 54: 24 para cada um dos peritos e de 6 para o sistema de ponderação.

Fez-se um estudo exaustivo para identificar os parâmetros da rede ME: efectuaram-se 30 corridas, onde em cada corrida o critério adoptado para terminar a optimização EM foi a técnica da validação cruzada. Seguidamente apresentam-se os resultados obtidos para os parâmetros identificados na corrida onde o erro de validação foi menor.

Caso 1: NH_4 como entrada no sistema de ponderação.

As Figuras 6.2 e 6.3 mostram as cinéticas identificadas em conjunto com as cinéticas 'experimentais' para a partição de treino e para a partição de validação respectivamente. Verifica-se que as taxas cinéticas 'experimentais' foram identificadas de forma aceitável tanto no treino como na validação.

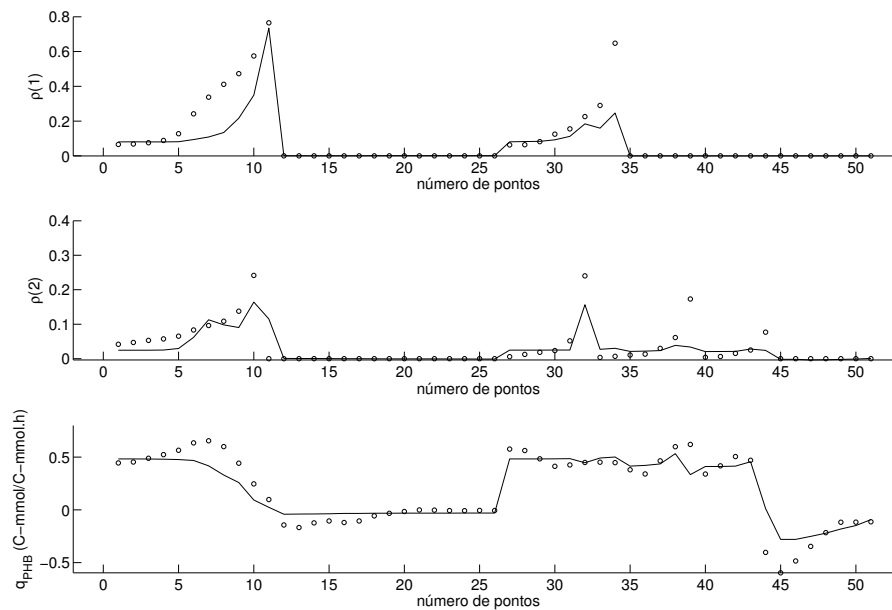


Figura 6.3: Resultados obtidos para $\rho(1)$, $\rho(2)$ e q_{PHB} para a partição de validação (caso1): valores experimentais (o, pontos), modelo ME (-, linha).

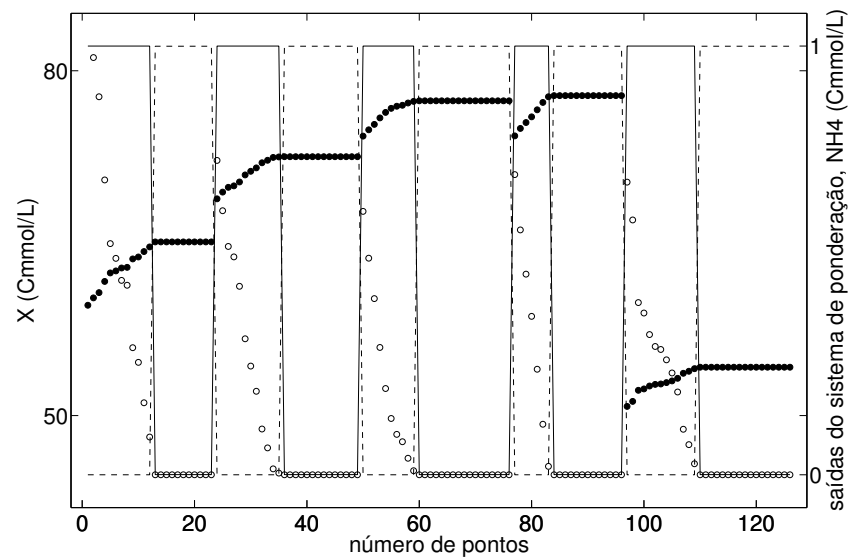


Figura 6.4: Saídas do sistema de ponderação localizado para a partição de treino (caso1): g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de NH_4 (o, pontos brancos) e X (●, pontos pretos)

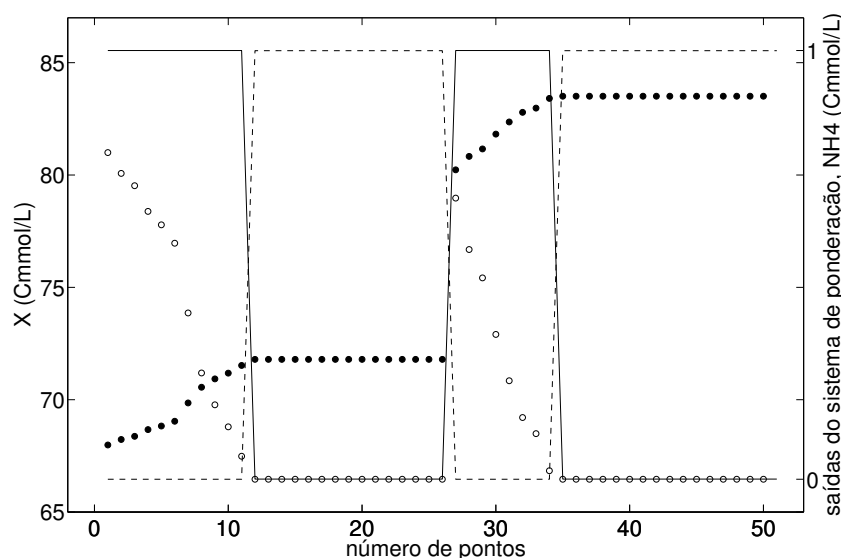


Figura 6.5: Saídas do sistema de ponderação localizado para a partição de validação (caso1): g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de NH_4 (o, pontos brancos) e X (●, pontos pretos)

A característica mais importante e mais interessante da rede ME foi observada mais uma vez neste caso de estudo como já tinha sido nos casos de estudo do Capítulo 5. A rede ME é capaz de detectar a transição entre a presença e a ausência de amónia como é ilustrado nas Figuras 6.4 e 6.5. Verifica-se que o modelo ME detecta a transição no ponto exacto onde a amónia passa de valores maiores que zero ($NH_4 > 0$) para valores iguais a zero ($NH_4 = 0$) tanto na partição de treino como na de validação. Os dois peritos especializam-se a modelizar as cinéticas de um ou de outro estado metabólico, neste caso, o estado de crescimento celular e o estado do não crescimento celular.

Os resultados obtidos com o modelo híbrido equações de balanço material/ME são apresentados na Figura 6.6 para a partição de treino e na Figura 6.7 para a partição de validação. Da análise das figuras conclui-se que o modelo híbrido consegue modelizar com uma exactidão aceitável os perfis das concentrações da biomassa, X , do ácido acético, HAc , da amónia, NH_4 e da fracção de PHB , f_{PHB} .

O erro absoluto médio de validação obtido foi de 0.21 C-mmol/L para a biomassa, de 6.27 C-mmol/L para o ácido acético, de 0.042 N-mmol/L para a amónia e de 0.14 C-mmol/C-mmol para a fracção de PHB . Para a partição de treino obteve-se um erro absoluto médio de 0.27 C-mmol/L para a biomassa, de 4.057 C-mmol/L para o ácido acético, de 0.053 N-mmol/L para a amónia e de 0.082 C-mmol/C-mmol para a fracção de PHB .

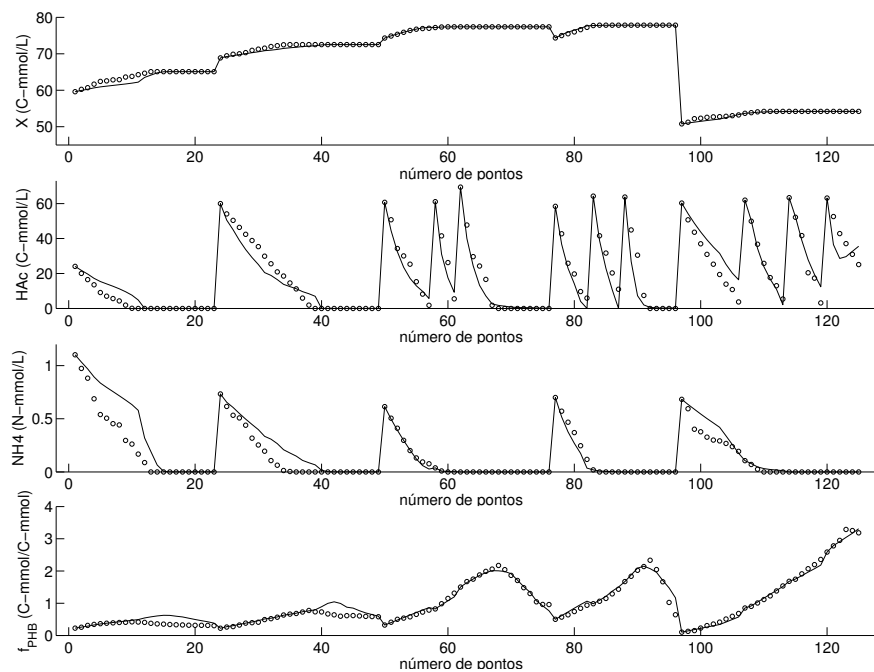


Figura 6.6: Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de treino (caso1): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).

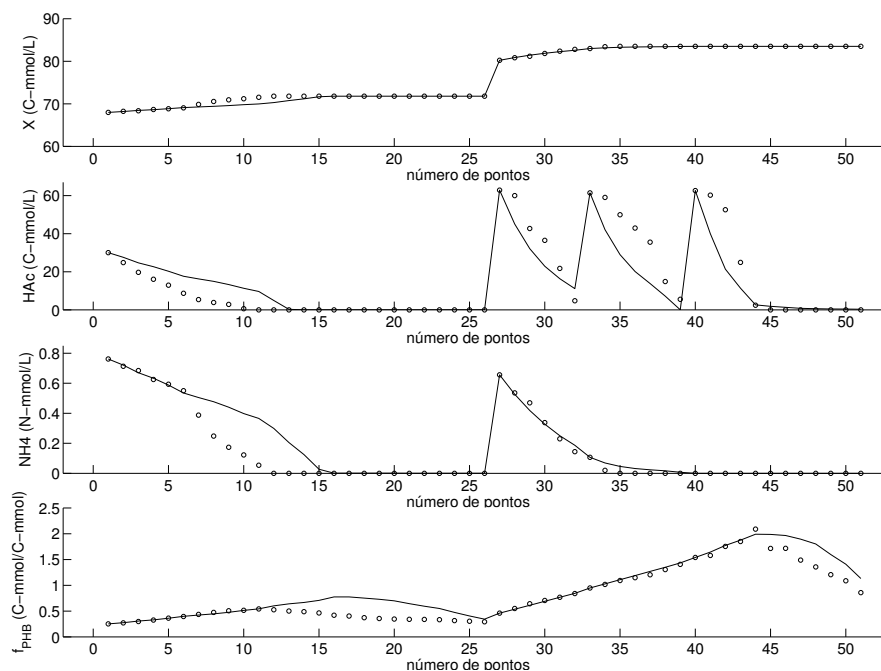


Figura 6.7: Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de validação (caso1): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).

Caso 2: HAc como entrada no sistema de ponderação.

As Figuras 6.8 e 6.9 mostram as cinéticas identificadas em conjunto com as cinéticas 'experimentais' para a partição de treino e para a partição de validação, respectivamente.

Verifica-se que as taxas cinéticas identificadas acompanham de forma razoável as cinéticas 'experimentais' tanto na partição de treino, como na partição de validação. Os erros aparentam ser superiores para $\rho(1)$, em que o modelo não consegue descrever alguns pontos da taxa de reacção.

Mais uma vez se verifica que a rede ME é capaz de detectar a transição entre a presença e a ausência de ácido acético como é ilustrado nas Figuras 6.10 e 6.11. Verifica-se que o modelo ME detecta a transição no ponto exacto onde o ácido acético passa de valores maiores que zero ($HAc > 0$) para valores iguais a zero ($HAc = 0$) tanto na partição de treino como na de validação. Os dois peritos especializam-se a modelizar as cinéticas de um ou de outro estado metabólico, neste caso, a fase de 'fome' e a fase da 'fartura'.

Os resultados obtidos com o modelo híbrido equações de balanço material/ME são apresentados na Figura 6.12 para a partição de treino e na Figura 6.13 para a partição de validação.

Da análise das figuras conclui-se que o modelo híbrido consegue modelizar com uma exactidão aceitável os perfis das concentrações da biomassa, X , do ácido acético, HAc , da amónia, NH_4 e da fracção de PHB , f_{PHB} .

O erro absoluto médio de validação obtido foi de 0.19 C-mmol/L para a biomassa, de 4.44 C-mmol/L para o ácido acético, de 0.037 N-mmol/L para a amónia e de 0.089 C-mmol/C-mmol para a fracção de PHB . Para a partição de treino obteve-se um erro absoluto médio de 0.18 C-mmol/L para a biomassa, de 4.044 C-mmol/L para o ácido acético, de 0.035 N-mmol/L para a amónia e de 0.064 C-mmol/C-mmol para a fracção de PHB . Os erros para a partição de validação são, portanto, ligeiramente superiores aos de treino.

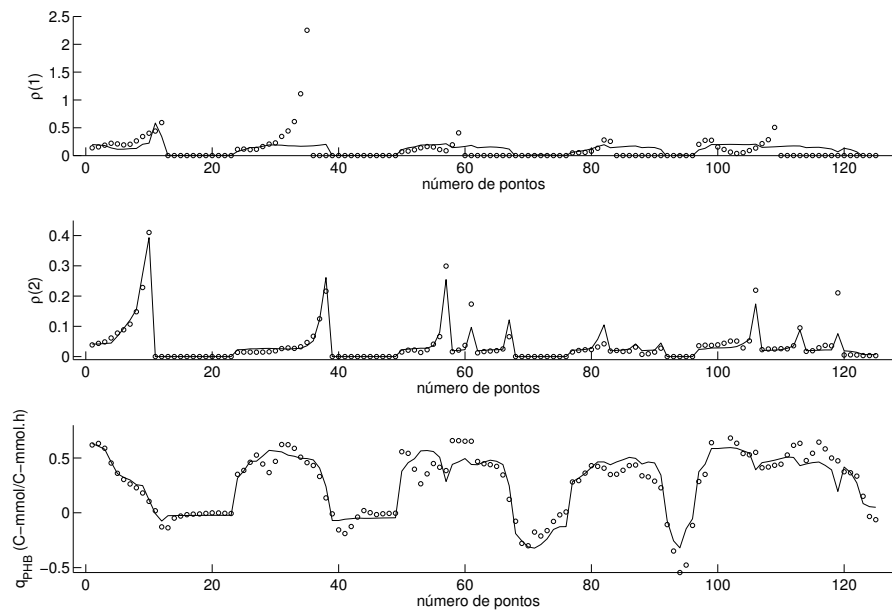


Figura 6.8: Resultados obtidos para $\rho(1)$, $\rho(2)$ e q_{PHB} para a partição de treino (caso2): valores experimentais (o, pontos), modelo ME (-, linha).

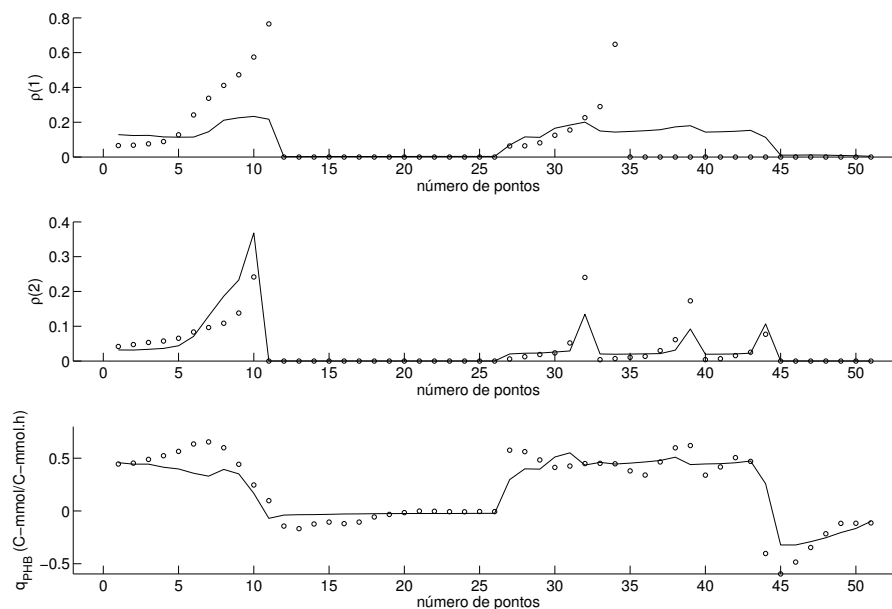


Figura 6.9: Resultados obtidos para $\rho(1)$, $\rho(2)$ e q_{PHB} para a partição de validação (caso2): valores experimentais (o, pontos), modelo ME (-, linha).

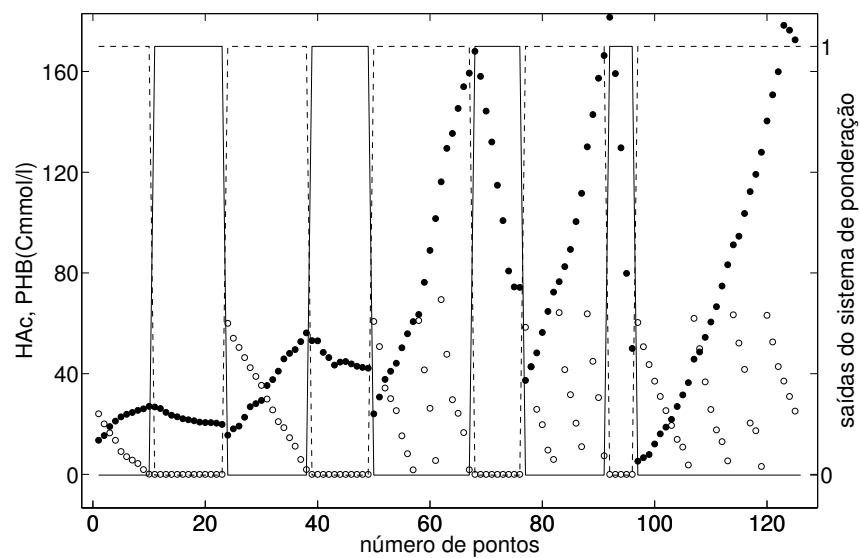


Figura 6.10: Saídas do sistema de ponderação localizado para a partição de treino (caso2): g_1 (—, linha sólida), g_2 (---, linha a tracejado) versus concentrações de HAc (o, pontos brancos) e PHB (●, pontos pretos)

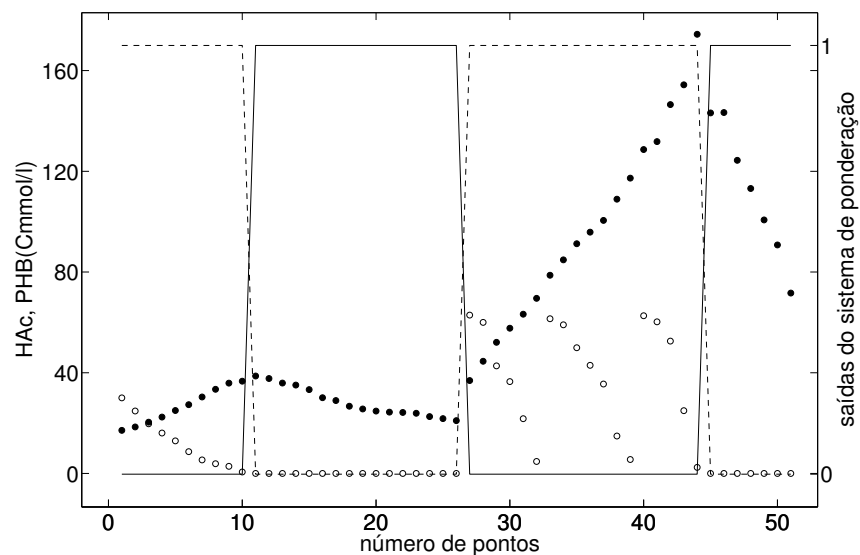


Figura 6.11: Saídas do sistema de ponderação localizado para a partição de validação (caso2): g_1 (—, linha sólida), g_2 (---, linha a tracejado) versus concentrações de HAc (o, pontos brancos) e PHB (●, pontos pretos)

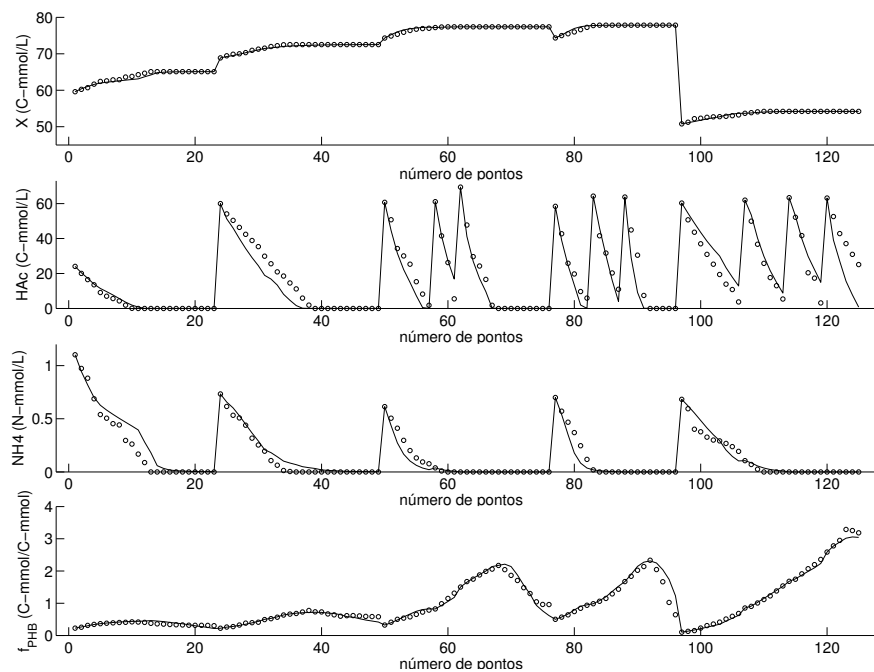


Figura 6.12: Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de treino (caso2): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).

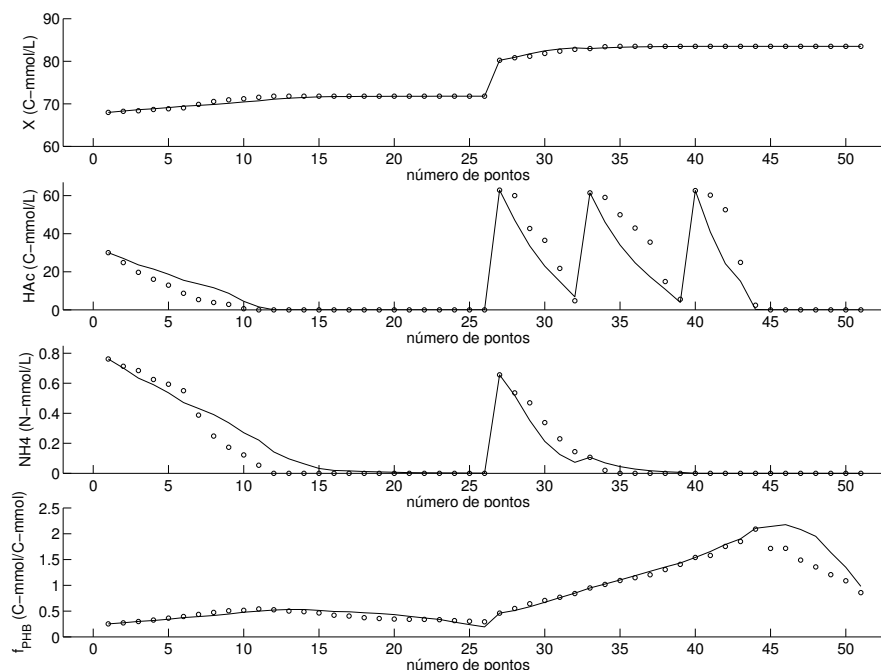


Figura 6.13: Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de validação (caso2): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).

6.3.2 Estratégia II: minimização dos erros nas concentrações

Caso 1: NH_4 como entrada no sistema de ponderação.

A rede ANN escolhida para integrar cada perito esquematizado na Figura 6.1(b) é do tipo MLP com 2 nodos na camada interna. As entradas da rede são HAc , NH_4 e f_{PHB} e as saídas são $\rho(1)$, $\rho(2)$ e q_{PHB} . Como o sistema de ponderação só tem uma variável de entrada, o modelo híbrido 'equações de balanço material'/ME para modelizar o processo de produção de PHB perfaz um total de 40 parâmetros: 17 parâmetros para cada perito e 6 para o sistema de ponderação.

Os parâmetros são identificados aplicando o algoritmo EM descrito na Secção 6.2.3. A Figura 6.14 apresenta o erro de treino e o erro de validação obtidos. Os parâmetros seleccionados para o modelo correspondem aos parâmetros obtidos na iteração cujo erro de validação é mínimo. Neste caso corresponde à iteração 15.

O erro absoluto médio de validação obtido foi de 0.16 C-mmol/L para a biomassa, de 3.70 C-mmol/L para o ácido acético, de 0.031 N-mmol/L para a amónia e de 0.086 C-mmol/C-mmol para a fracção de PHB . Para a partição de treino obteve-se um erro absoluto médio de 0.11 C-mmol/L para a biomassa, de 2.96 C-mmol/L para o ácido acético, de 0.023 N-mmol/L para a amónia e de 0.0898 C-mmol/C-mmol para a fracção de PHB . Como seria de esperar, o modelo é um pouco mais exacto na descrição da partição de treino do que na partição de validação.

Os resultados da modelização para a partição de validação e para a partição de treino são apresentados nas Figuras 6.15 e 6.16, respectivamente. Da análise das figuras conclui-se que o modelo híbrido consegue modelizar com grande exactidão os perfis das concentrações da biomassa, X , do ácido acético, HAc , da amónia, NH_4 e de f_{PHB} .

Como mostram as Figuras 6.17 e 6.18 o modelo híbrido é capaz de detectar a transição entre o estado de crescimento celular e o estado de não crescimento celular, coincidente com a presença e a ausência de amónia. Os dois peritos especializam-se a modelizar as cinéticas de um ou de outro estado metabólico.

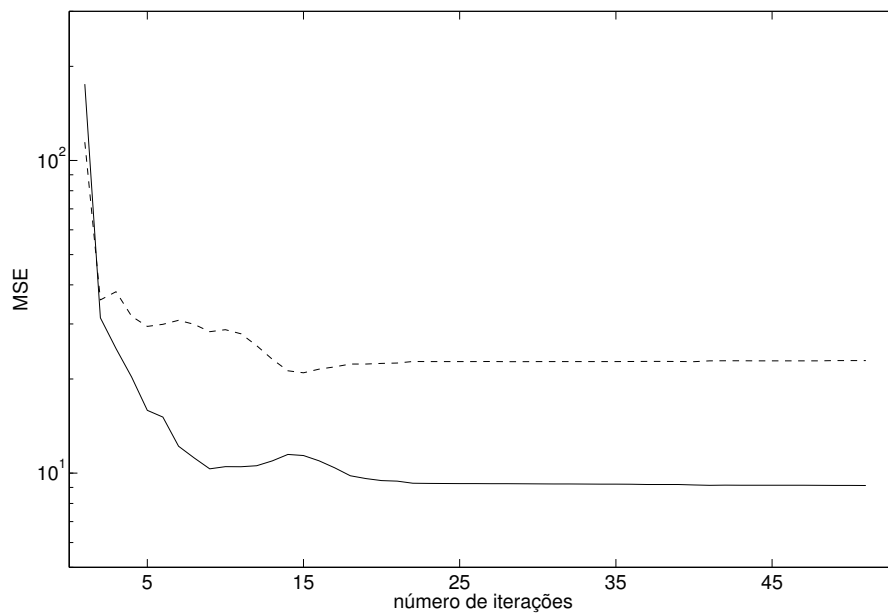


Figura 6.14: Erro quadrático médio (MSE) em função do número de iterações (caso1): para a partição de treino (—, linha a cheio) e para a partição de validação (---, linha a tracejado).

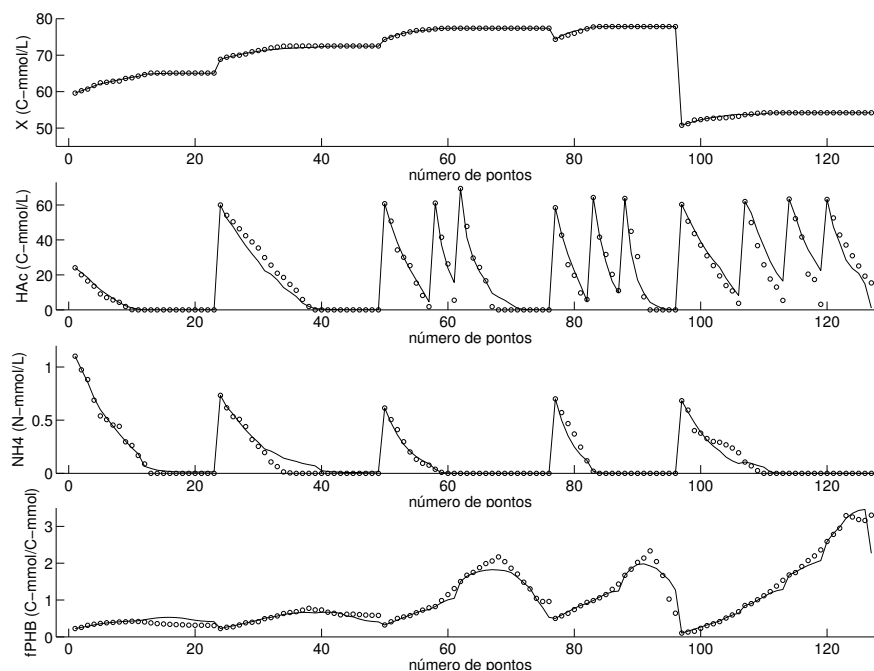


Figura 6.15: Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de treino (caso1): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (—, linha).

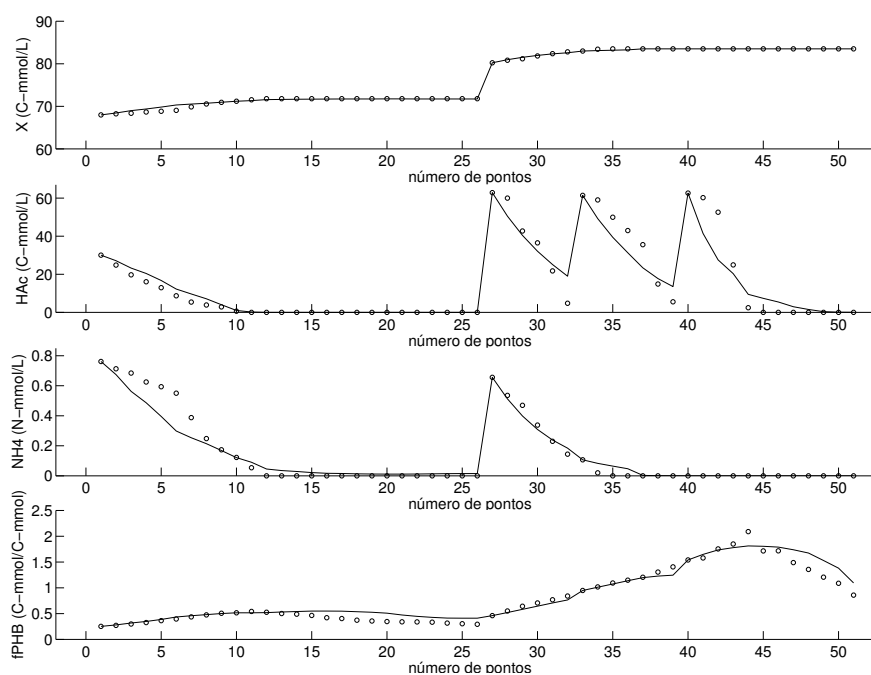


Figura 6.16: Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de validação (caso1): valores experimentais (o, pontos), modelo híbrido ME/equações de balanço de massa (-, linha).

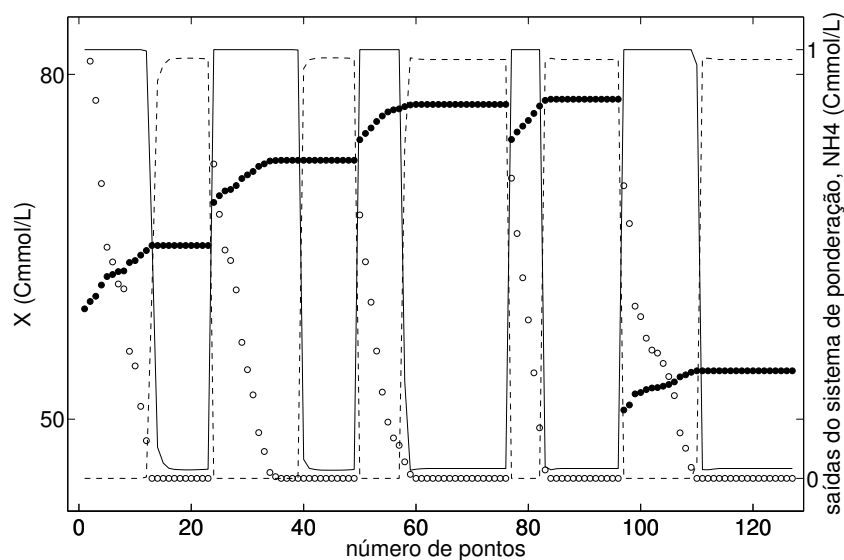


Figura 6.17: Saídas do sistema de ponderação localizado para a partição de treino (caso1): g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de NH_4 (o, pontos brancos) e X (•, pontos pretos)

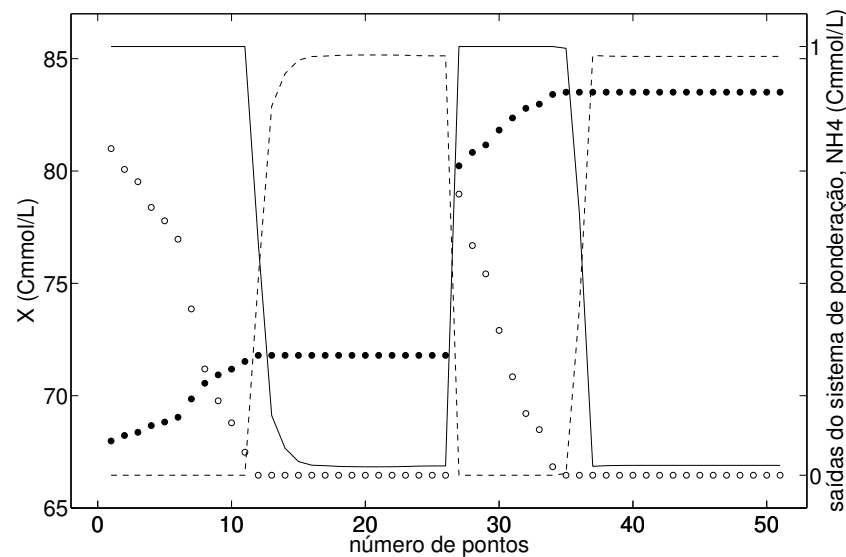


Figura 6.18: Saídas do sistema de ponderação localizado para a partição de validação (caso1): g_1 (—, linha sólida), g_2 (---, linha a tracejado) versus concentrações de NH_4 (o, pontos brancos) e X (•, pontos pretos)

Caso 2: HAc como entrada no sistema de ponderação.

A rede ANN escolhida para integrar cada perito esquematizado na Figura 6.1(b) é do tipo MLP com 3 nodos na camada interna. As entradas da rede são HAc , NH_4 e f_{PHB} e as saídas são $\rho(1)$, $\rho(2)$ e q_{PHB} , como no caso 1. Como o sistema de ponderação só tem uma variável de entrada, o modelo híbrido 'equações de balanço material'/ME para modelizar o processo de produção de PHB perfaz um total de 54 parâmetros: 24 parâmetros para cada perito e 6 para o sistema de ponderação.

Os parâmetros são identificados aplicando o algoritmo EM descrito na Secção 6.2.3. A Figura 6.19 apresenta o erro de treino e o erro de validação obtido. Os parâmetros seleccionados para o modelo correspondem aos parâmetros obtidos na iteração cujo erro de validação é mínimo. Neste caso corresponde à iteração 99.

O erro absoluto médio obtido na partição de validação foi de 0.16 C-mmol/L para a biomassa, de 4.56 C-mmol/L para o ácido acético, de 0.033 N-mmol/L para a amónia e de 0.063 C-mmol/C-mmol para a fracção de PHB . O erro absoluto médio obtido no treino foi de 0.092 C-mmol/L para a biomassa, de 3.41 C-mmol/L para o ácido acético, de 0.018 N-mmol/L para a amónia e de 0.091 C-mmol/C-mmol para a fracção de PHB .

Os resultados da modelização para a partição de validação e para a partição de treino são apresentados nas Figuras 6.20 e 6.21, respectivamente. Da análise das figuras

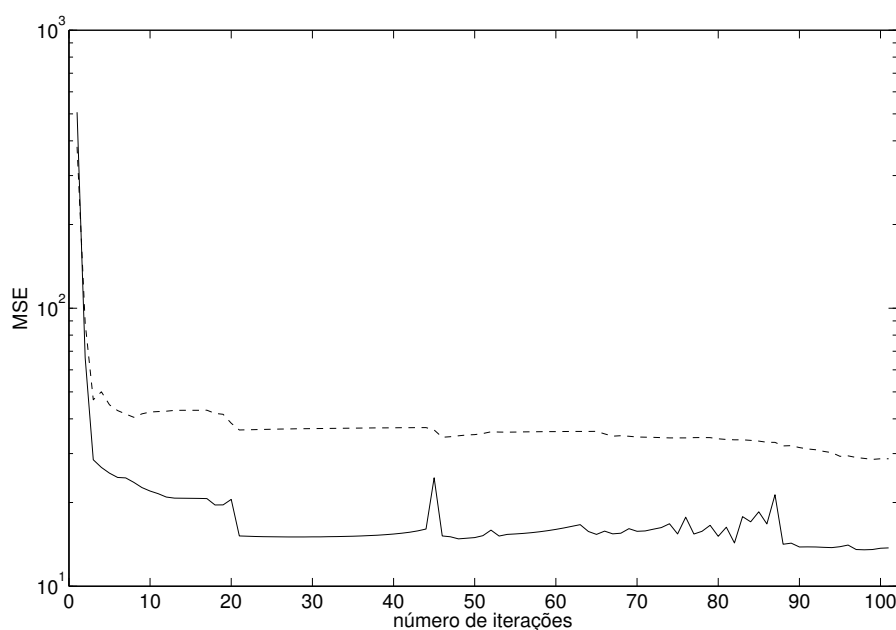


Figura 6.19: Erro quadrático médio (MSE) em função do número de iterações para a partição de treino (caso2): (—, linha a cheio) e para a partição de validação (---, linha a tracejado).

Tabela 6.1: Erros de validação absolutos médios para as duas estratégias de identificação, para os vários casos estudados

	Estratégia I				Estratégia II			
	X	HAC	NH_4	f_{PHB}	X	HAC	NH_4	f_{PHB}
	C-mmol/L	C-mmol/L	N-mmol/L	C-mmol/C-mmol	C-mmol/L	C-mmol/L	N-mmol/L	C-mmol/C-mmol
Caso 1	0.21	6.27	0.042	0.14	0.16	3.70	0.031	0.086
Caso 2	0.19	4.44	0.037	0.089	0.16	4.56	0.033	0.063

conclui-se que o modelo híbrido, analogamente ao caso anterior, consegue modelizar com grande exactidão os perfis das concentrações da biomassa, X , do ácido acético, HAC , da amónia, NH_4 e de f_{PHB} . Também neste caso, a exactidão é superior na descrição da partição de treino.

Como mostram as Figuras 6.22 e 6.23 o modelo híbrido é capaz de detectar a transição entre a fase de 'fome' e de 'fartura', i.e., detecta a presença e a ausência de ácido acético. Os dois peritos especializam-se a modelizar as cinéticas de um ou de outro estado metabólico. A Tabela 6.1 resume os resultados obtidos para as duas estratégias de identificação de parâmetros em função dos casos estudados.

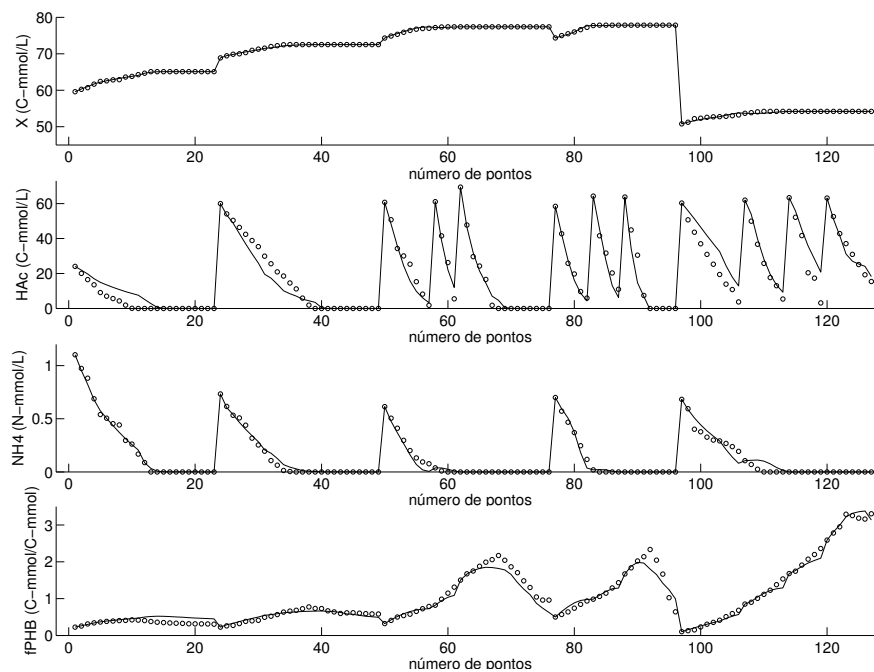


Figura 6.20: Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de treino (caso2): valores experimentais (o, pontos), modelo híbrido ME/eqações de balanço de massa (-, linha).

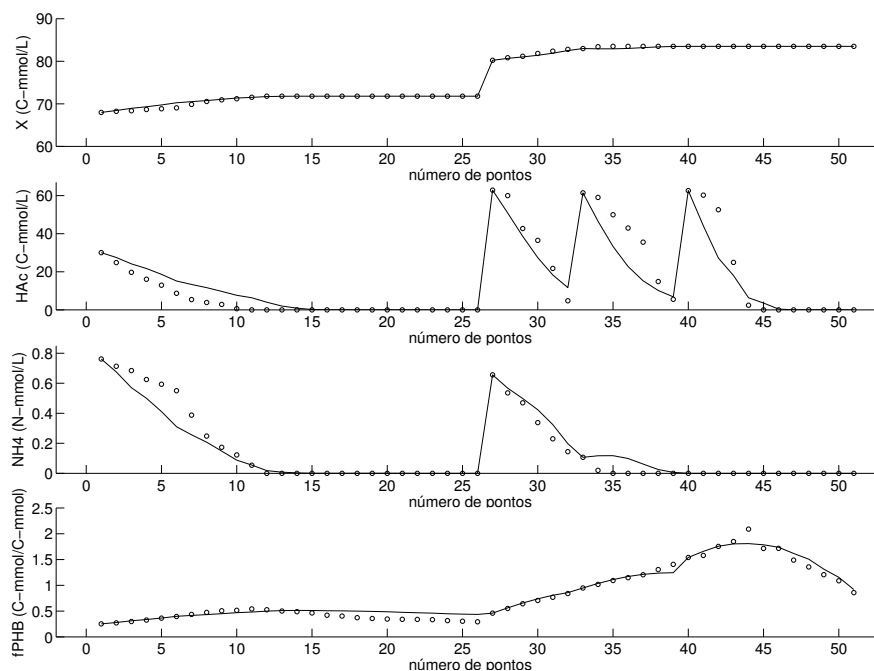


Figura 6.21: Resultados obtidos para X , HAc , NH_4 e f_{PHB} para a partição de validação (caso2): valores experimentais (o, pontos), modelo híbrido ME/eqações de balanço de massa (-, linha).

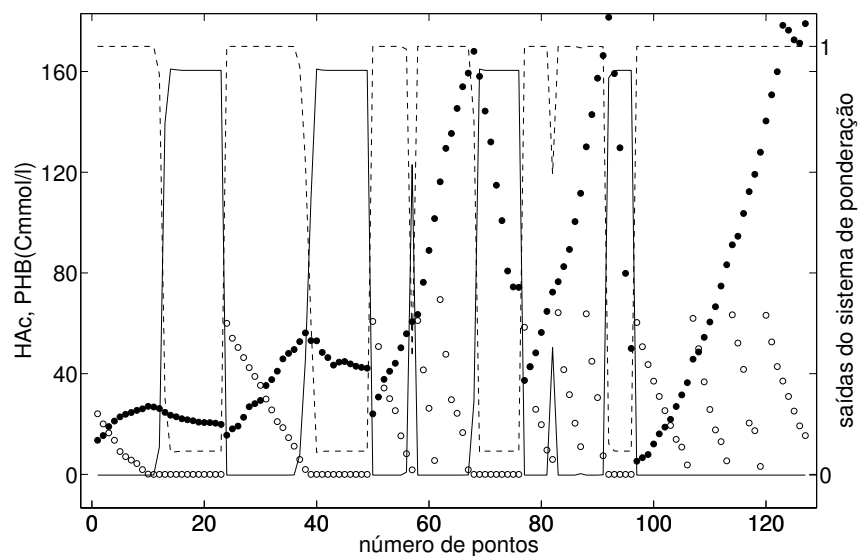


Figura 6.22: Resultados para a partição de treino (caso2): saídas do sistema de ponderação localizado: g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de HAc (o, pontos brancos) e PHB (●, pontos pretos)

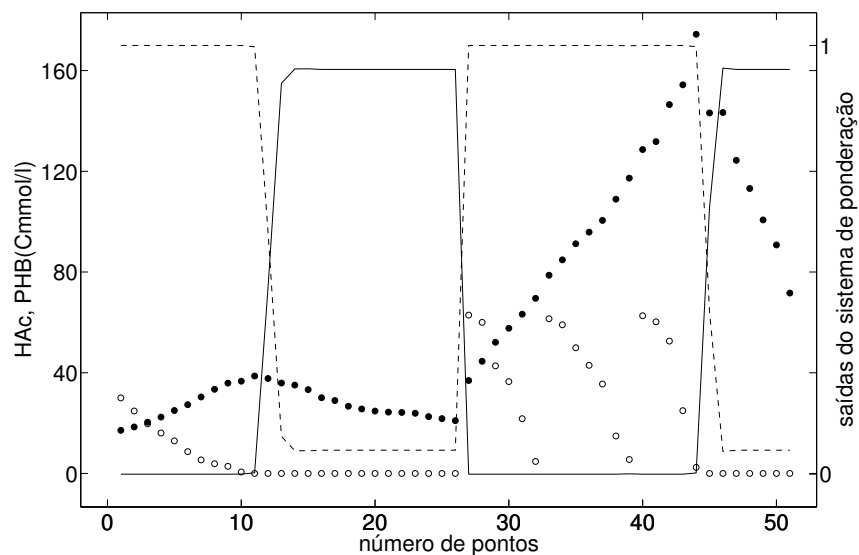


Figura 6.23: Resultados para a partição de validação (caso2): saídas do sistema de ponderação localizado: g_1 (-, linha sólida), g_2 (- -, linha a tracejado) versus concentrações de HAc (o, pontos brancos) e PHB (●, pontos pretos)

6.4 Conclusões

Neste capítulo estudou-se uma estrutura híbrida semelhante à do Capítulo 4, mas onde as descrições das cinéticas é feita por redes de mistura de peritos.

Como se mostrou no Capítulo 5, o sistema 'cinética celular' possui algumas particularidades nomeadamente uma estrutura intrínseca modular a qual pode ser modelizada com vantagem usando redes modulares de peritos.

O treino destas redes juntamente com balanços materiais é agora mais complexo. O algoritmo EM foi adaptado a este novo modelo híbrido e as duas estratégias de identificação de parâmetros descritas no Capítulo 4 foram comparadas.

Estas metodologias foram usadas para modelizar um processo de produção de Poli-hidroxicanoatos à escala laboratorial e foram validadas pela técnica da validação cruzada.

Foram estudados dois casos onde ocorre uma transição metabólica: o caso onde a transição ocorre entre o estado de crescimento celular e não crescimento celular e o caso onde a transição ocorre entre a fase de 'fome' e a fase de 'fartura'.

Demonstrou-se que o modelo híbrido balanço material/ME foi capaz de detectar estas transições e de se organizar em módulos que têm correspondência com os estados metabólicos do ciclo do processo.

Conclui-se que se obtiveram melhores resultados, i.e., um modelo mais exacto, quando se aplica a estratégia de identificação baseada na minimização dos erros nas concentrações no caso considerado no estudo. Com efeito, no Capítulo 4 obtiveram-se as mesmas conclusões.

É importante realçar que, em cada uma das experiências apresentadas, as células são sujeitas a uma fase de 'fome' com a duração de 9 horas. Esta fase provoca uma adaptação fisiológica a condições externas de limitação de carbono intracelular a qual se reflecte numa variabilidade experiência a experiência. Este comportamento dinâmico do 'sistema célula' dificulta a obtenção de um modelo exacto. Mesmo assim, os resultados exibem exactidão suficiente para posteriores estudos de optimização dinâmica. A suportar esta conclusão, vem o facto de a partição de validação ter sido sempre descrita com exactidão comparável à da partição de treino. Perspectiva-se portanto, que esta técnica de modelização possa constituir uma ferramenta atractiva para optimização de processos com culturas mistas sujeitas a condições transientes de alimentação, tais como os processos de tratamento de águas residuais.

Referências

Henze, M., Gujer, W., Mino, T. e van Loosdrecht, M. E. (Editores) (2000), *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3; Scientific and Technical Report 9*, IWA Publishing, London.

Capítulo 7

Mistura Híbrida de Peritos Baseados em Diferentes Paradigmas de Modelização

Conteúdo do Capítulo

No Capítulo 2 foram revistos sumariamente os métodos de modelização híbrida existentes na literatura. Neste capítulo é proposta uma nova estratégia de conjugar os diferentes tipos de conhecimento, para a modelização de processos químicos e bioquímicos: conhecimento mecanístico, conhecimento heurístico e conhecimento escondido em registos de dados do processo, baseada em mistura de peritos. O conceito de Rede Modular Baseada em Conhecimento (*Knowledge Based Modular network - rede KBM*) é apresentado. O algoritmo da Esperança-Maximização (EM) é empregue para combinar de forma óptima os peritos dentro da estrutura de rede KBM. Os conceitos são ilustrados com a aplicação da produção de fermento de padeiro. Os resultados da identificação mostram que é possível obter uma descrição mais exacta do processo quando todas as fontes de conhecimento disponível acerca do processo são incorporadas no modelo do mesmo.

7.1 Introdução

A modelização pode ser vista como um exercício de expressão e representação do conhecimento disponível numa forma compacta. Uma vez que numa maneira geral existem e estão disponíveis diferentes tipos de conhecimento sobre o mesmo processo, duas abordagens conceptuais para modelizar são essencialmente possíveis:

1. seleccionar uma das fontes de conhecimento disponíveis e adoptar e explorar a correspondente técnica de modelização; ou
2. incorporar e expressar todas as fontes disponíveis de conhecimento num único modelo híbrido do processo.

Na perspectiva da engenharia do conhecimento a segunda alternativa é mais vantajosa uma vez que a exactidão do modelo depende da 'quantidade' e da 'qualidade' do conhecimento disponível. A aplicação desta segunda alternativa resulta em estruturas mais complexas que podem ser classificadas em série ou em paralelo ou simultaneamente em série e paralelo como as estruturas apresentadas no Capítulo 6. Em estruturas paralelas ocorre competição entre os submodelos, sendo necessário um sistema de ponderação.

No Capítulo 2 faz-se referência aos métodos de ponderação mais importantes descritos na literatura para o tipo de estruturas competitivas (ou paralelas). Concretamente, são referidos dois métodos: métodos de ponderação baseados em técnicas de agrupamento e métodos de ponderação baseados em sistemas difusos. No entanto, estes métodos apresentam algumas desvantagens nomeadamente, serem direccionados para resolverem um problema específico e serem baseados na confiança heurística da capacidade extrapolativa dos diferentes modelos que expressam os diferentes tipos de conhecimento existentes. A questão essencial é que nenhum destes métodos faz a avaliação da exactidão de cada modelo em regiões diferentes do espaço das entradas e além disso, nenhum deles inclui essa informação na estratégia de ponderação.

De facto, se o objectivo é modelizar um processo de forma tão exacta quanto possível, é fundamental que a exactidão de cada modelo para cada entrada seja considerada no método de ponderação. Consequentemente, neste capítulo é proposta uma estrutura computacional para modelização híbrida genérica de processos baseada em Mistura de Peritos que toma em consideração esta questão. São aplicados e desenvolvidos os conceitos e os algoritmos das Redes de Mistura de Peritos (Jacobs e Jordan, 1991; Jacobs et al., 1991) para o campo da modelização híbrida baseada em diferentes paradigmas de modelização.

7.2 Uma nova estrutura híbrida: rede modular baseada em conhecimento

A linha mestra deste método consiste numa estrutura de rede conexionista com nós (peritos) especializados capazes de representar diferentes tipos de conhecimento com diferentes níveis de sofisticação.

A rede, designada por, Rede Modular Baseada em Conhecimento (*Knowledge Based Modular network* - rede KBM), é então capaz de incorporar e misturar diferentes tipos de conhecimento e técnicas de modelização. É dada a esta estrutura uma interpretação probabilística e a técnica usada para identificação dos parâmetros é baseada na máxima verosimilhança. Parâmetros, esses, existentes em cada um dos peritos e parâmetros que dizem respeito às ligações entre peritos.

Concretamente, Jacobs e Jordan (1991), Jacobs et al. (1991) e Jordan e Jacobs (1994) desenvolveram uma classe de arquiteturas modulares conexionista designadas por mistura de peritos (ME) e subsequentemente por mistura hierárquica de peritos (*Hierarchical mixtures of experts* - HME). A característica principal deste tipo de arquiteturas conexionistas é a de serem capazes de aprender a dividir uma tarefa em duas ou mais tarefas independentes e atribuir peritos distintos a aprender cada uma destas subtarefas.

Com efeito, estes mesmos conceitos podem ser aplicados para desenvolver uma rede que integra modelos de tipos diferentes que têm a capacidade de se tornarem especialistas na descrição das diferentes partes dum processo. Esta é a ideia que está subjacente à definição da Rede Modular Baseada em Conhecimento.

A rede KBM está representada esquematicamente na Figura 7.1. Esta rede consiste num conjunto de K peritos baseados em diferentes paradigmas de modelização mediado por um sistema de ponderação.

Concretamente, a tarefa da rede é associar o espaço de entrada S_{inp} ao espaço de saída D_{out} . Tanto o perito como o sistema de ponderação tem acesso às entradas $\mathbf{x} \in S_{inp}$, que podem ser iguais para todos os peritos mas não necessariamente. A tarefa de cada perito i é aproximar uma função $f_i : S_{inp} \rightarrow D_{out}$ sobre uma região do espaço de entrada S_{inp} .

A tarefa do sistema de ponderação é atribuir um perito a cada vector de entrada \mathbf{x} . A saída final \mathbf{y} é a combinação linear das saídas dos peritos \mathbf{y}_i pesados pelas saídas g_i do sistema de ponderação. Os detalhes dos peritos e do sistema de ponderação são

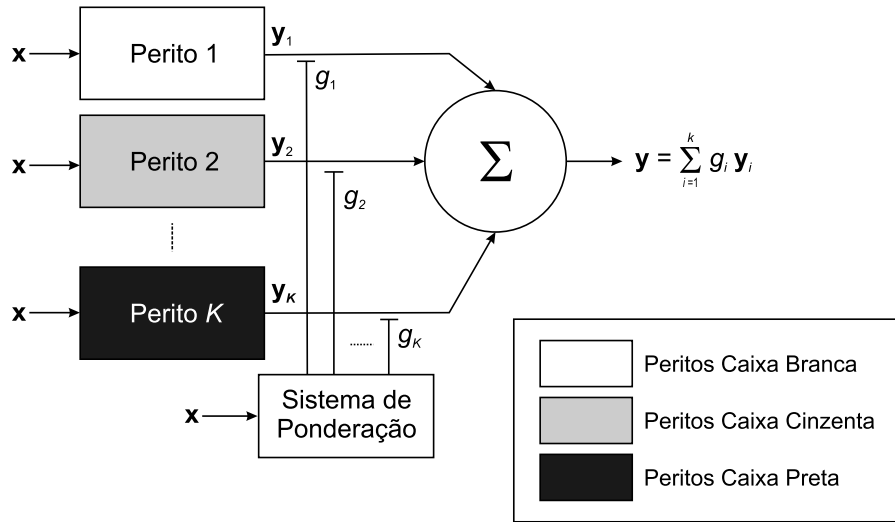


Figura 7.1: Diagrama de blocos da rede KBM composta por K peritos. As saídas de cada perito são mediadas por um sistema de ponderação. Os diferentes peritos expressam diferentes tipos de conhecimento: de caixa branca, de caixa cinzenta e de caixa preta

descritos nas secções 7.2.1 e 7.2.2.

7.2.1 Definição dos Peritos

Jacobs e Jordan (1991) e Jordan e Jacobs (1994) utilizaram peritos lineares. Todavia, Weigend et al. (1995) exploraram o uso de peritos não lineares e designaram a arquitectura desenvolvida por *gated mixture of experts*. No caso das estruturas de rede KBM os peritos deverão ser, ou é esperado que o sejam, mais elaborados no sentido em que devem expressar diferentes formas de conhecimento.

Neste trabalho os peritos são formalizados da seguinte forma: designando por \mathbf{x} o vector de entrada e por \mathbf{d} o vector de saída, assume-se que os padrões medidos $\{\mathbf{x}, \mathbf{d}\} \in \{S_{inp}, D_{out}\}$ são gerados por um conjunto de K processos regressivos diferentes contínuos não lineares e dinâmicos por natureza. Assim sendo, os peritos são definidos da seguinte forma:

$$\frac{d\mathbf{y}_i}{dt} = f_i(\mathbf{y}_i, \mathbf{x}, \mathbf{w}_i), \quad (7.1)$$

em que \mathbf{y}_i é o vector de saída do perito i , \mathbf{x} o vector de entrada, \mathbf{w}_i o vector de parâmetros do perito i e f_i uma função contínua não linear.

As relações funcionais f_i não são as mesmas para todos os peritos. Daí que, se assume que os peritos podem incluir conhecimento disponível de diferentes formas e níveis de

complexidade teórica. Posto isto, os peritos podem ser de três tipos:

1. peritos de caixa branca que expressam o conhecimento físico por intermédio de equações matemáticas;
2. peritos de caixa cinzenta que englobam vários tipos de conhecimento, nomeadamente, mecanístico, heurístico ou baseado em dados do processo;
3. peritos de caixa preta que são capazes de extrair informação a partir dos dados do processo (por exemplo, ANNs)

Pressupõe-se que cada forma de conhecimento representará melhor o processo numa determinada região do espaço de entrada. O papel da rede KBM é dividir o espaço de entrada em sub-regiões de tal forma que o conhecimento expresso é maximizado através da especialização dos peritos em cada sub-região.

Com esta estratégia poder-se-á possivelmente extrair o melhor que cada tipo de conhecimento tem para oferecer. Esta característica pode constituir uma vantagem decisiva relativamente aos métodos de ponderação referidos no Capítulo 2.

7.2.2 Definição do Sistema de Ponderação

Como já descrito nos Capítulos 5 e 6 existem dois tipos principais de sistemas de ponderação, para as estruturas de rede modular da Figura 7.1:

- (i) baseado em funções *softmax* (Jacobs e Jordan, 1991);
- (ii) baseado em funções Gaussianas (Jacobs et al., 1991; Ramamurti e Ghosh, 1999).

Outro aspecto importante, neste tipo de redes, que importa mencionar é a relação que existe entre a natureza dos peritos e a natureza do sistema de ponderação. Com peritos muito simples, por exemplo, lineares, a rede deverá ser configurada com muitos peritos de maneira a efectuar um mapeamento aceitável. No entanto, com peritos mais sofisticados serão necessários apenas alguns peritos para se chegar a um resultado satisfatório. Neste último caso o sistema de ponderação deve permitir uma divisão flexível do espaço de influência de cada perito. Nesta secção propõe-se uma alternativa aos sistemas (i) e (ii) por forma a obter-se partições flexíveis quando se usam peritos complexos.

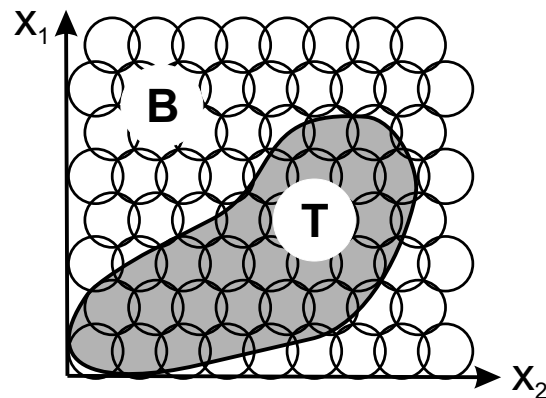


Figura 7.2: Sistema de ponderação: representação esquemática dum conjunto hipersférico de grupos distribuído equidistantemente num espaço de entradas bidimensional. A região cinzenta define o subespaço T onde existem medidas do processo em estudo.

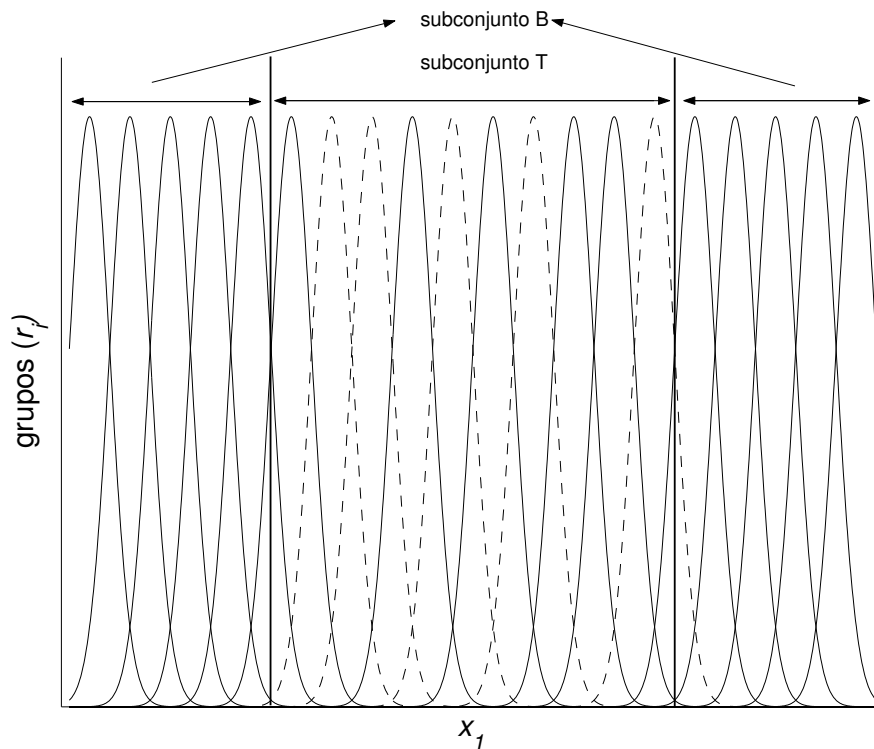


Figura 7.3: Sistema de ponderação: exemplo da associação entre grupos ($N_C = 20$) e peritos ($K = 2$) para o caso unidimensional do espaço das entradas. A linha a cheio representa o subconjunto $\{C\}_1$ de $N_T/2 + N_B$ grupos associados ao perito 1 ($\tau_j = 1$); a linha a tracejado representa o subconjunto $\{C\}_2$ de $N_T/2$ grupos associados ao perito 2 ($\tau_j = 2$).

Tendo estas questões em consideração propõe-se um sistema de ponderação formalizado em dois passos:

Passo 1: Agrupamento hiperesférico do espaço das entradas S_{inp} com um conjunto C de N_C grupos distribuídos equidistantemente ao longo do espaço das entradas S_{inp} . Os grupos são definidos como funções Gaussianas:

$$\nu(\mathbf{x}, \mathbf{m}_j, \mathbf{\Sigma}_j) = (2\pi)^{-n/2} |\mathbf{\Sigma}_j|^{-1/2} \exp\{(\mathbf{x} - \mathbf{m}_j)^T \mathbf{\Sigma}_j^{-1} (\mathbf{x} - \mathbf{m}_j)\} \quad (7.2)$$

sendo \mathbf{m}_j os centros do grupo e $\mathbf{\Sigma}_j = \text{diag}\{\sigma_{ij}^2\}$ a diagonal da matriz das covariâncias.

O conjunto C de grupos é dividido em dois subconjuntos: o subconjunto T de N_T grupos dentro da região onde existem medidas do processo; e o subconjunto B de N_B grupos fora desta região. A Figura 7.2 ilustra este conceito, mostrando a projecção das regiões para o caso de duas entradas x_1 e x_2 .

Passo 2: Associação de cada perito a um subconjunto de grupos $\{C\}_i (i = 1, \dots, K)$.

Dentro da região de treino T , N_T/K grupos são associados aleatoriamente a um perito específico i . Adicionalmente, todos os N_B grupos na região B são associados com o perito mecanístico com melhores propriedades de extrapolação.

Para o universo de N_C grupos, a associação entre grupos e peritos é formalizada através dum vector de parâmetros inteiros τ , definido da seguinte forma:

$$\tau_j = i \text{ para } j = 1, \dots, N_C \quad (7.3)$$

onde $\tau_j = i$ significa que o grupo j é associado ao perito i .

Exemplo (ilustrado na Figura 7.3): uma entrada x_1 ; dois peritos ($K = 2$), um total de 20 grupos ($N_C = 20$); subconjunto B com 10 grupos ($N_B = 10$); subconjunto T com 10 grupos ($N_T = 10$); perito 1 (mecanístico) associado a 15 grupos ($N_T/2 = 5$ grupos escolhidos aleatoriamente de T mais todos os grupos de B); perito 2 associado a 5 grupos ($N_T/2 = 5$ grupos escolhidos aleatoriamente de T).

O cálculo do peso relativo g_i de cada perito i está directamente relacionado com as funções do grupo, definidas pela Equação (7.2), sendo obtido da seguinte forma:

1. dados os padrões de entrada \mathbf{x} e dado o vector de parâmetros inteiros τ , para cada perito i escolhe-se o grupo mais próximo de centro θ_i e correspondente matriz das covariâncias $\mathbf{\Lambda}_i$ tal que $\tau_j = \tau$,

2. cálculo dos grupos ν_i usando a Equação (7.2):

$$\nu_i = \nu(\mathbf{x}, \theta_i, \mathbf{\Lambda}_i), \quad i = 1, \dots, K \quad (7.4)$$

3. cálculo das saídas normalizadas g_i do sistema de ponderação a partir da equação seguinte:

$$g_i = \frac{\nu_i}{\sum_{l=1}^K \nu_l} \quad i = 1, \dots, K. \quad (7.5)$$

7.2.3 Identificação de parâmetros

Pode ser dada uma interpretação probabilística à estrutura de rede KBM e o método da Esperança-Maximização pode ser utilizado para a identificação dos parâmetros. A interpretação probabilística é baseada na presunção de que o processo que gera os padrões \mathbf{d} a partir dos padrões \mathbf{x} obedece a determinado modelo probabilístico $P(\mathbf{d}|\mathbf{x})$.

Para um dado perito i assume-se que a probabilidade condicional do padrão \mathbf{d} (valor medido) de dimensão n_d , dado o vector de entrada \mathbf{x} de dimensão n_x e dado o perito i , é representada por uma função de probabilidade Gaussiana:

$$P(\mathbf{d}|\mathbf{x}, \mathbf{w}_i) = 2\pi^{-n_d/2} |\Sigma_i|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{d} - \mathbf{y}_i)^T \Sigma_i^{-1} (\mathbf{d} - \mathbf{y}_i) \right\}, \quad (7.6)$$

em que Σ_i é a matriz covariância para o perito i e \mathbf{y}_i o vector de saída do perito i que determina o centro da função Gaussiana. Note-se que \mathbf{y}_i é função dos parâmetros \mathbf{w}_i como definido na equação 7.1.

As saídas do sistema de ponderação também são interpretadas como a probabilidade condicional de escolher o perito i dado o vector de entradas \mathbf{x} :

$$P(i|\mathbf{x}) = g_i(\mathbf{x}, \tau), \quad (7.7)$$

Com esta interpretação probabilística é possível formular a função de verosimilhança e empregar o estimador de máxima verosimilhança para simultaneamente estimar o vector de parâmetros \mathbf{w}_i de cada perito e o vector de parâmetros τ do sistema de ponderação.

Jacobs et al. (1991), para maximizar a função da verosimilhança, aplicaram para o caso mais simples de mistura de peritos o algoritmo de actualização dos pesos pelos gradientes ascendentes. Mais tarde, Jordan e Jacobs (1994) sugeriram o uso do algo-

ritmo da Esperança-Maximização (EM) que provaram convergir mais rapidamente que o algoritmo dos gradientes ascendentes.

Adoptou-se o algoritmo EM que consiste num procedimento iterativo a dois passos (Jordan e Jacobs, 1994; Xu e Jordan, 1996) onde cada iteração p pode ser resumida da seguinte forma:

1. o passo-E, onde as probabilidades posteriores h_i são calculadas a partir de:

$$h_i^{p,t} = \frac{g_i(\mathbf{x}^t, \tau^p) P(\mathbf{d}^t | \mathbf{x}^t, \mathbf{w}_i^p)}{\sum_{j=1}^K g_j(\mathbf{x}^t, \tau^p) P(\mathbf{d}^t | \mathbf{x}^t, \mathbf{w}_j^p)} \quad t = 1, \dots, n_p, \quad (7.8)$$

em que o sobrescrito t refere-se a um padrão medido e n_p representa o número total de padrões medidos.

2. o passo-M, onde $K + 1$ problemas de maximização são resolvidos:

$$\mathbf{w}_i^{p+1} = \operatorname{argmax} \left(\sum_{t=1}^{n_p} h_i^{p,t} \ln P(\mathbf{d}^t | \mathbf{x}^t, \mathbf{w}_i^p) \right) \quad i = 1, \dots, K \quad (7.9)$$

$$\tau^{p+1} = \operatorname{argmax} \left(\sum_{t=1}^{n_p} \sum_{j=1}^K h_j^{p,t} \ln g_j^t(\mathbf{x}^t, \tau^p) \right) \quad (7.10)$$

As K otimizações da Equação (7.9) calculam os novos parâmetros \mathbf{w}_i do perito, por sua vez a otimização da Equação (7.10) calcula os novos parâmetros τ para o sistema de ponderação (isto corresponde a uma reassociação 'grupo-para-perito').

O algoritmo pára quando as reassociações entre grupos e peritos deixam de ocorrer entre duas iterações seguidas.

7.3 Caso de Estudo II: processo de produção do fermento de padeiro

7.3.1 Formulação do problema

Neste caso de estudo ilustra-se a aplicação duma rede KBM simples para prever a evolução da concentração da biomassa ao longo do tempo como função do perfil da taxa de alimentação de fonte de carbono num processo semicontínuo de produção de

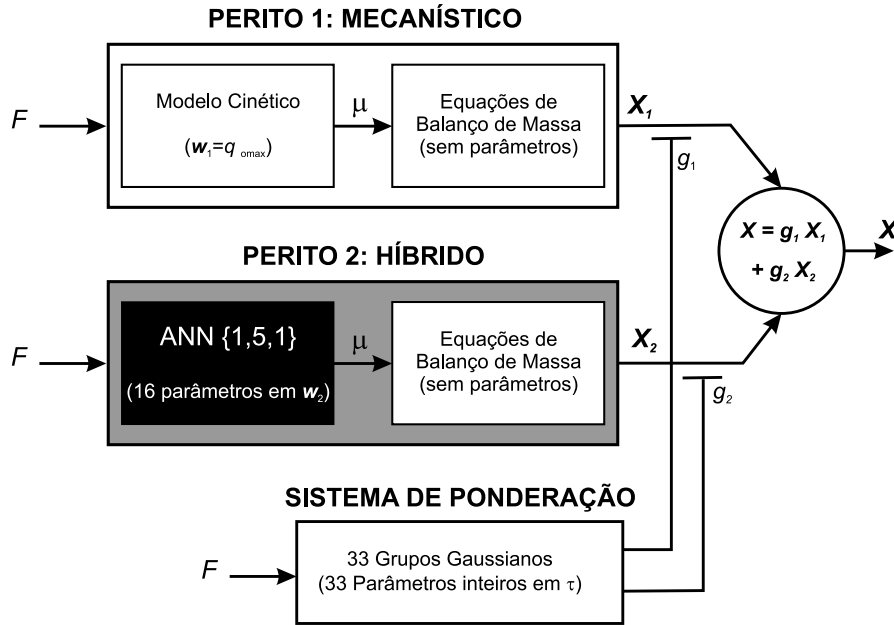


Figura 7.4: Rede KBM para a predição da biomassa num processo de fermentação do fermento de padeiro com $K = 2$ peritos. O perito 1 é um modelo mecanístico (caixa branca) e o perito 2 é um modelo híbrido (caixa cinzenta)

fermento de padeiro. Dados de cinco fermentações (Oliveira, 1997), B1, B2, B3, B4 e B7 descritas na Secção 3.2.1 do Capítulo 3, foram utilizados para desenvolver a rede KBM.

Pomerleau e Perrier (1990) propôs um modelo dinâmico para o processo semicontínuo do fermento de padeiro que é baseado na descrição mecanística do processo. Neste caso de estudo o objectivo principal é combinar este modelo mecanístico com outro modelo que incorpora outras fontes de conhecimento. Sendo assim, duas principais fontes de conhecimento são utilizadas:

1. Conhecimento mecanístico: Considera-se que o conhecimento mecanístico disponível é expresso apropriadamente pelo modelo mecanístico de Pomerleau e Perrier (1990);
2. Informação escondida nos dados do processo: considera-se que características importantes do processo não incorporadas no modelo mecanístico estão reflectidas no conjunto de medidas disponíveis.

A estratégia é a de activar as duas fontes de conhecimento na rede KBM para prever a concentração da biomassa. A rede KBM é composta por dois peritos mediados por um sistema de ponderação de acordo com a Figura 7.4.

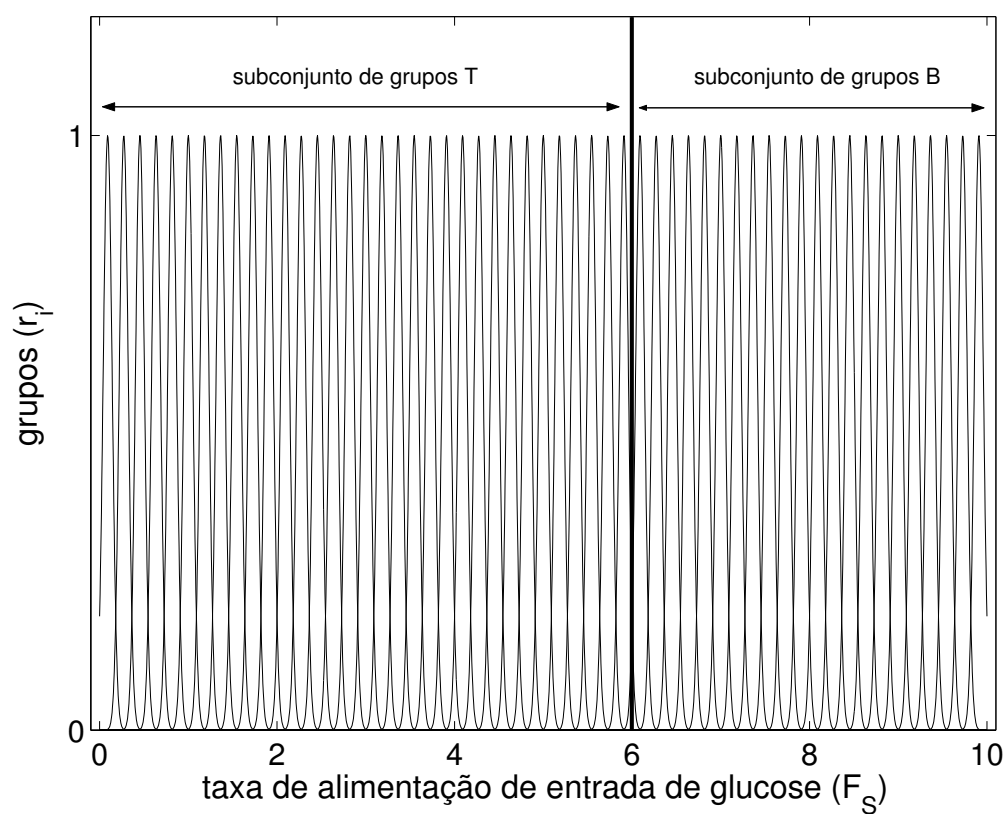


Figura 7.5: Distribuição dos grupos no sistema de ponderação da rede KBM para o fermento de padeiro. Número total de grupos é $N_C = 55$; $N_T = 33$ grupos pertencendo ao subconjunto T e $N_B = 22$ grupos pertencendo ao subconjunto B

O perito 1 é o modelo mecanístico dinâmico descrito na Secção 3.2.1 do Capítulo 3, definido pelas Equações (3.16)-(3.21) do mesmo capítulo. O parâmetro q_{omax} (taxa máxima de consumo de oxigénio) é um dos parâmetros do modelo mais sensíveis e é uma característica de uma dada estirpe (Sonnleitner e Kappeli, 1986). Por razões de simplicidade, considera-se que o parâmetro q_{omax} é o único parâmetro que precisa de ser identificado para as condições actuais do processo.

Assim, o perito 1 envolve só um parâmetro na sua identificação que é o q_{omax} .

O perito 2 é um modelo híbrido simples (Figura 7.4) baseado na equação de balanço de massa para a biomassa X_2 dada pela equação:

$$\frac{dX_2}{dt} = \left(\mu(F_S) - \frac{F}{V} \right) X_2 \quad (7.11)$$

em que μ é a taxa específica de crescimento, V o volume de trabalho e F_S a taxa de alimentação de glucose por unidade de volume definida pela expressão:

$$F_S = F \frac{S_o}{V} \quad (7.12)$$

em que S_o é a concentração de glucose na corrente F . Considera-se que a taxa específica de crescimento é desconhecida e é uma função não linear de F_S . A ideia é empregar uma ANN para aproximar a função $\mu(F_S)$.

Foi escolhida uma rede MLP de pequena dimensão com três camadas e com funções de activação sigmóides (ver Capítulo 4). A rede tem uma entrada, cinco nodos internos e uma saída. Esta estrutura de rede corresponde a um total de $\mathbf{w}_2 = 16$ parâmetros.

O sistema de ponderação foi configurado com 55 grupos Gaussianos distribuídos equidistantemente no intervalo $F_S \in [0, 10]$ g/(L.h) com desvios padrão de $\sigma = 0.05$ g/(L.h). O subconjunto T de grupos é composto por 33 grupos no subintervalo $F_S \in [0, 6]$ g/(L.h). O subconjunto B de grupos é composto por 22 grupos no subintervalo complementar de $F_S \in [0, 6]$ g/(L.h). Esta configuração é ilustrada na Figura 7.5.

O vector de parâmetros τ é inicializado aleatoriamente de forma a associar o perito 1 e o perito 2 com 16 e 17 grupos do subconjunto T , respectivamente. Adicionalmente, todos os 22 grupos do subconjunto B foram associados ao perito 1. Um total de 55 parâmetros inteiros estão envolvidos no sistema de ponderação.

A rede KBM tem, assim, 17 parâmetros reais e 55 parâmetros inteiros que têm de ser identificados empregando o algoritmo EM (Equações (7.8), (7.9) e (7.10)).

As duas optimizações referentes à Equação (7.9) foram efectuadas empregando um algoritmo quasi-Newton (da biblioteca *Numerical Algorithms Group* - NAG). A optimização da Equação (7.10) foi efectuada empregando um algoritmo de recozimento simulado (*Simulated annealing algorithm*) (Cardoso et al., 1994).

7.4 Resultados da identificação

Os resultados da identificação produzidos pelo algoritmo EM são bastante sensíveis aos valores iniciais dos parâmetros. Se, inicialmente, um perito descrever melhor, i.e., de modo mais exacto, a dinâmica do processo em todo o espaço das entradas, o algoritmo EM não converge para a combinação óptima dos dois peritos. O resultado será que só um perito é usado para o mapeamento. Para obstar a este problema, num passo inicial, adaptou-se independentemente os dois peritos às medidas experimentais.

Os resultados da predição da biomassa, depois deste primeiro passo, estão indicados juntamente com as medidas experimentais na Figura 7.6(a)-(e). A estatística usada para comparar os modelos é o erro quadrático médio (MSE) definido da seguinte forma:

$$MSE = \frac{1}{P} \sum_{t=1}^P (X_{m,t} - X_t)^2 \quad (7.13)$$

sendo $X_{m,t}$ a biomassa medida.

Os gráficos das Figuras 7.6(a)-(e) mostram que, qualitativamente, ambos os peritos predizem razoavelmente bem os valores da biomassa de todas as fermentações, com a excepção das fermentações (B2) e (B4) onde as predições do perito mecanístico são particularmente más.

O MSE inicial foi de 15.00 e 1.28 para os peritos mecanístico e híbrido, respectivamente. Uma análise qualitativa dos gráficos mostra que para alguns pontos medidos, o perito mecanístico dá melhores predições que o perito híbrido. Este é um bom ponto de partida para aplicar o algoritmo EM.

Num segundo passo, o algoritmo EM foi empregue para a identificação dos parâmetros. Os resultados obtidos após 10 iterações são apresentados nas Figuras 7.6(a)-(f). O MSE final foi de 0.35, o que representa uma melhoria significativa relativamente às predições iniciais de ambos os peritos.

Apenas 3 de 33 grupos do subconjunto T foram atribuídos ao perito mecanístico.

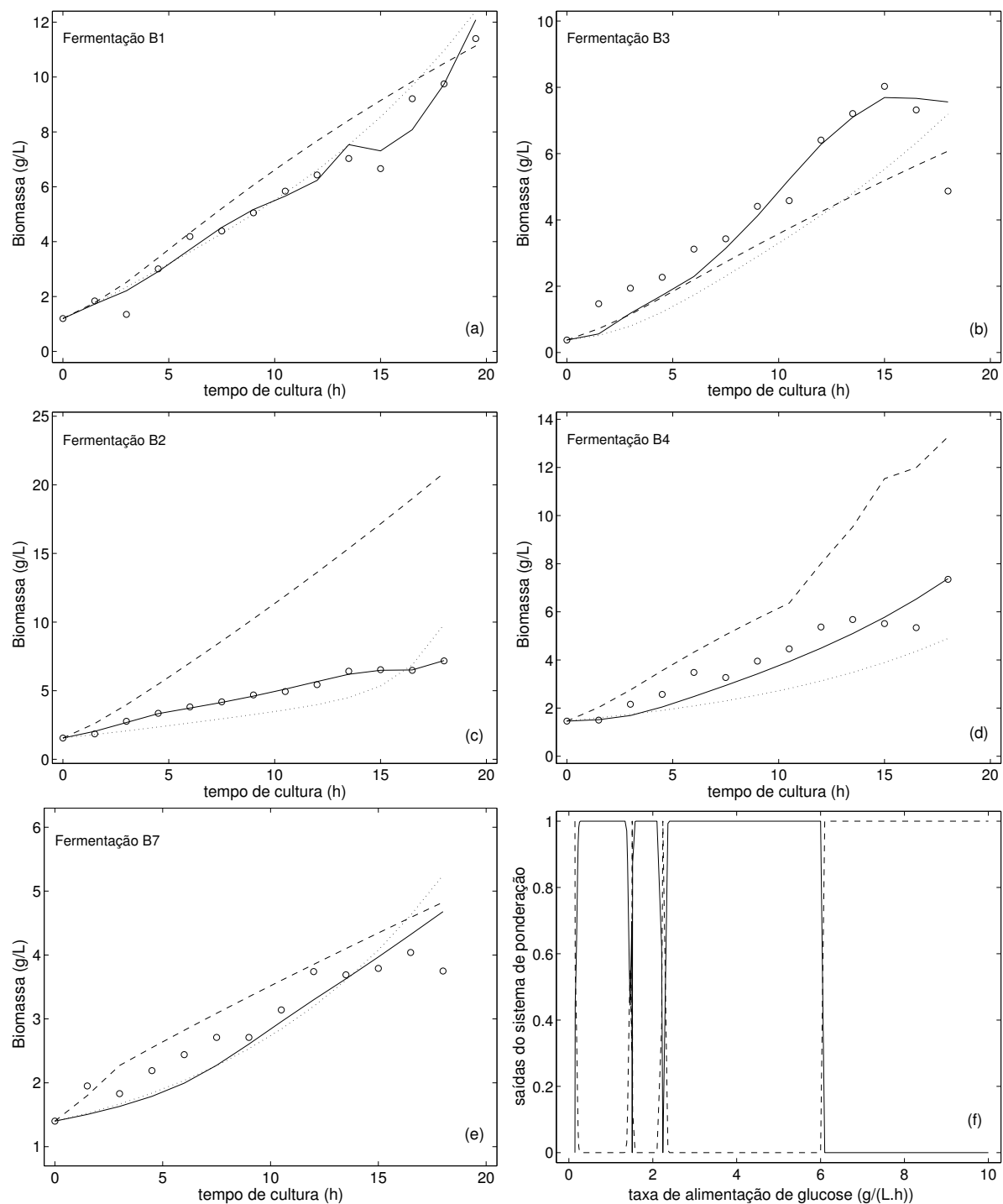


Figura 7.6: Curvas de predição da biomassa como função do tempo de cultura num processo semicontínuo de fermentação do fermento de padreiro.(a)-(e) Resultados para 5 fermentações. Os círculos representam os valores de biomassa medidos. As linhas a pontilhado e tracejado representam os resultados da predição dos peritos 1 e 2, respectivamente, antes de aplicar o algoritmo EM. As linhas a cheio representam os resultados da predição pela rede KBM. (f) Saídas do sistema de ponderação em função da taxa de alimentação da glucose, g_1 (linha a tracejado) referente ao perito 1, g_2 (linha a cheio) referente ao perito 2

Isto quer dizer que o perito híbrido é capaz de prever de forma mais exacta o conjunto de medidas da biomassa do que o perito mecanístico.

Na Figura 7.6(f) as saídas do sistema de ponderação são apresentadas em função da taxa de alimentação de glucose. Na gama de $F_S \in [0, 6]$ g/(L.h), a saída do sistema de ponderação correspondente ao perito híbrido (g_2) é quase sempre 1 ao passo que g_1 é 0.

Na gama de $F_S \in [6, 10]$ g/(L.h) apenas o perito mecanístico é usado. Este facto está de acordo com as especificações iniciais dos grupos do subconjunto B serem atribuídos à partida ao perito mecanístico, o qual em princípio será mais fiável em condições de extrapolação do que outras técnicas de modelização.

7.5 Conclusões

Neste capítulo propôs-se um método novo para melhorar a qualidade da modelização de processos através da integração de conhecimento de diferentes níveis de sofisticação. O conceito principal é a estrutura designada por Rede Modular Baseada em Conhecimento (rede KBM) que objectivamente fornece uma organização geral para combinar vários tipos de conhecimento usualmente disponíveis tanto nos processos químicos como bioquímicos: o mecanístico, o heurístico e o conhecimento escondido nos dados do processo. O método procura uma combinação óptima entre os diversos modelos/fontes de conhecimento utilizando métodos estatísticos bem conhecidos. Este método tem a grande vantagem de entrar em consideração com a exactidão de cada modelo em diferentes regiões do espaço das entradas. Assim, estamos a garantir que em cada instante o modelo mais exacto é usado para calcular a saída final do modelo.

O conceito foi ilustrado e testado com uma aplicação de produção de fermento de padeiro à escala laboratorial. Definiram-se dois peritos que envolvem paradigmas de modelização diferentes, nomeadamente, um dos peritos é um modelo mecanístico e o outro é um modelo híbrido do tipo equações de balanço material/ANN. Estes peritos foram treinados dentro do domínio experimental da operação. O uso combinado destes peritos numa rede de mistura de peritos levou a melhoramentos na capacidade de predição comparada com os desempenhos individuais de cada perito.

Uma possível dificuldade que pode obstar à propagação deste tipo de abordagem será a sua própria complexidade. Contudo, com a capacidade computacional existente nos computadores modernos esta dificuldade poderá ser facilmente ultrapassada construindo

programas de computador dedicados, flexíveis e amigáveis.

Ao aumentar a eficiência da informação utilizada está-se a aperfeiçoar a exactidão dos modelos. Note-se, que em optimização de processos, a exactidão do modelo é uma questão crítica. Quanto mais exacto for o modelo mais fiáveis serão as estratégias óptimas de operação calculadas.

Referências

- Cardoso, M. F., Salcedo, R. L. e Deazevedo, S. F. (1994), 'Nonequilibrium simulated annealing a faster approach to combinatorial minimization', *Industrial & Engineering Chemistry Research*, **33**(8), 1908–1918.
- Jacobs, R. A. e Jordan, M. I. (1991), 'A competitive modular connectionist architecture', em J. M. R.P. Lippman e D. Touretzky (Editores), *Advances in Neural Information Processing Systems*, CA Morgan Kaufmann, San Mateo, volume 3, Páginas 767–773.
- Jacobs, R. A., Jordan, M. I. e Barto, A. G. (1991), 'Task decomposition through competition in a modular connectionist architecture the what and where vision tasks', *Cognitive Science*, **15**(2), 219–250.
- Jordan, M. I. e Jacobs, R. A. (1994), 'Hierarchical mixtures of experts and the em algorithm', *Neural Computation*, **6**(2), 181–214.
- Oliveira, F. M. (1997), *Monitorização e Controlo de Fermentadores: Aplicação ao Fermento de Padeiro*, Tese de Doutoramento, Faculdade de Engenharia da Universidade do Porto.
- Pomerleau, Y. e Perrier, M. (1990), 'Estimation of multiple specific growth-rates in bioprocesses', *AIChE Journal*, **36**(2), 207–215.
- Ramamurti, V. e Ghosh, J. (1999), 'Structurally adaptive modular networks for nonstationary environments', *IEEE Transactions on Neural Networks*, **10**(1), 152–160.
- Sonnleitner, B. e Kappeli, O. (1986), 'Growth of *Saccharomyces-cerevisiae* is controlled by its limited respiratory capacity formulation and verification of a hypothesis', *Biotechnology and Bioengineering*, **28**(6), 927–937.
- Weigend, A. S., Mangeas, M. e Srivastava, A. N. (1995), 'Nonlinear gated experts for time series: Discovering regimes and avoiding overfitting', *International Journal of Neural Systems*, **6**(4), 373–399.

Xu, L. e Jordan, M. I. (1996), 'On convergence properties of the em algorithm for gaussian mixtures', *Neural Computation*, **8**(1), 129–151.

Capítulo 8

Conclusões e trabalho futuro

As técnicas de optimização e controlo tem vindo a tornar-se cada vez mais importante na indústria de processos devido ao aumento de competitividade do mercado. No entanto, a aceitação e a implementação de metodologias baseadas em modelos matemáticos para o melhoramento dos processos produtivos permanece reduzida, principalmente porque a razão benefícios/custos não é suficientemente atractiva para tais desenvolvimentos.

Os processos bioquímicos são complexos e normalmente pouco conhecidos do ponto de vista mecanístico, em particular no que diz respeito ao sistema 'população celular'. No método clássico modelos matemáticos são utilizados para descrever os mecanismos conhecidos. Devido à falta de conhecimentos mecanísticos e devido à elevada complexidade do sistema celular, modelos não-estruturados e não-segregados são frequentemente a única escolha possível. Enquanto que em processos estacionários estes modelos mais simples permitem por vezes resultados suficientemente exactos, em processos descontínuos e semicontínuos altamente dinâmicos, a simplificação em termos de estrutura e heterogeneidade celular resulta geralmente em modelos pouco exactos e com capacidade preditiva reduzida. Como consequência, a aceitação de modelos matemáticos para optimizar processos bioquímicos permanece bastante reduzida na indústria.

Em contrapartida, é reconhecido que a modelização híbrida do tipo caixa cinzenta é uma alternativa com custos de desenvolvimento inferiores, relativamente aos modelos matemáticos, para a análise de bioprocessos (Schubert et al., 1994b,a; Preusting et al., 1996; Simutis et al., 1997; van Can et al., 1998; Peres et al., 2001). O princípio basilar da concepção deste tipo de modelos é o de não considerar o conhecimento mecanístico como a única fonte de conhecimento relevante, considerando outras fontes de conhecimento - heurísticas ou informação escondida em bases de dados - como recursos complementares

importantes para o desenvolvimento de modelos. Para sistemas complexos, para os quais há falta de conhecimento mecanístico, este ponto é particularmente importante. Uma vez que os modelos híbridos integram mais conhecimento que os modelos clássicos baseados em princípios fundamentais, pode-se obter maior exactidão com menos experiências e portanto com custos de desenvolvimento inferiores.

A aplicação da modelização híbrida a reactores químicos e bioquímicos tem sido demonstrada em diversos trabalhos. A estrutura híbrida mais utilizada é baseada nas equações de balanço material, como na abordagem tradicional baseada em princípios fundamentais, no entanto as cinéticas são modelizadas por redes neuronais artificiais (Psichogios e Ungar, 1992; Thompson e Kramer, 1994; Montague e Morris, 1994; Feyerherm et al., 1997; van Can et al., 1998, 1999; Chen et al., 2000). As redes de percepção de camada múltipla (*Multiple Layer Perceptron* - MLP) e as redes de funções de base radial (*Radial Basis Function* - RBF) são as mais frequentes por duas razões: primeiro, porque foi provado que estas redes são funções não lineares de aproximação universais, e segundo, a sua aplicação não requer o conhecimento acerca da estrutura do sistema que se pretende modelizar.

Estes dois argumentos motivaram a sua aplicação para modelizar os mecanismos muito complexos associados com o crescimento celular e biocatálise (Montague e Morris, 1994). A combinação de ANNs com equações de balanço material e de energia em paralelo (Thompson e Kramer, 1994) ou em série (Psichogios e Ungar, 1992) constitui-se em sistemas não lineares dinâmicos descritos por um conjunto de equações diferenciais ordinárias (*Ordinary Differential Equations* - ODEs).

Na perspectiva da identificação e análise de sistemas dinâmicos existem muitas questões teóricas tais como a identifiabilidade e a estabilidade que não estão bem estudadas mesmo para as estruturas híbridas mais simples ANN/equações de balanço. A maior parte dos estudos referidos são dirigidos a um problema específico. De facto, na literatura não existem muitas referências a aplicações de modelos híbridos do tipo caixa cinzenta em controlo de processos devido, talvez, à falta de uma teoria sólida.

O principal objectivo desta tese é o desenvolvimento de técnicas de modelização híbrida do tipo caixa cinzenta, particularmente adequadas para processos biotecnológicos tendo a preocupação da sistematização destas técnicas tornando-as passíveis de ser utilizadas independentemente do processo em estudo.

Os tópicos investigados foram os seguintes:

1. caracterização teórica de estruturas híbridos do tipo caixa cinzenta;

2. identificação de parâmetros;
3. aplicabilidade a processos biotecnológicos.

De seguida apresentam-se as principais conclusões:

8.1 Estrutura de modelo híbrido

A descrição da dinâmica dum biorreactor perfeitamente agitado e com mistura completa é obtida duma maneira geral por um conjunto de equações de balanço material aos componentes que intervêm no mecanismo de reacção. As equações de balanço material constituem o conhecimento mecanístico básico que pode ser aplicado na maior parte dos casos com um grau de confiança elevado. Relativamente às taxas cinéticas não se pode dizer o mesmo, e o que se fez foi separar o termo cinético mecanístico conhecido de outro termo que pode ser modelizado por técnicas alternativas. Portanto, o modelo híbrido do tipo caixa cinzenta combina modelos mecanísticos com modelos não paramétricos: o sistema biorreactor é descrito por um conjunto de equações de balanço material e o sistema célula é representado por uma mistura flexível de representações não paramétricas e mecanísticas.

A Figura 8.1 mostra as estruturas de modelo híbrido estudadas neste trabalho. A estrutura apresentada e estudada no Capítulo 4 está esquematizada na Figura 8.1 como Estrutura I. Neste caso, o sistema célula é representado pela mistura ajustável dum termo cinético mecanístico conhecido com um termo cinético desconhecido que é modelizado por um modelo não paramétrico. O sistema biorreactor é representado pelo conjunto de equações de balanço material.

Este tipo de estrutura híbrida geral incorpora três propriedades importantes que contribuem para a sistematização do estudo da identificação e da estabilidade permitindo elaborar estratégias avançadas de optimização e controlo, a saber:

1. a representação das cinéticas através da mistura de um modelo mecanístico com um modelo não paramétrico é bastante genérica tanto para reacções de catálise químicas como biológicas (p.e., Bastin e Dochain 1990; Dochain et al. 1992).
2. a organização introduzida por esta expressão permite o uso de outras técnicas de modelização para calcular ρ . Em vez de uma só rede neuronal, m redes neuronais, um sistema difuso ou uma série de combinações destas estruturas ou doutras são possíveis.

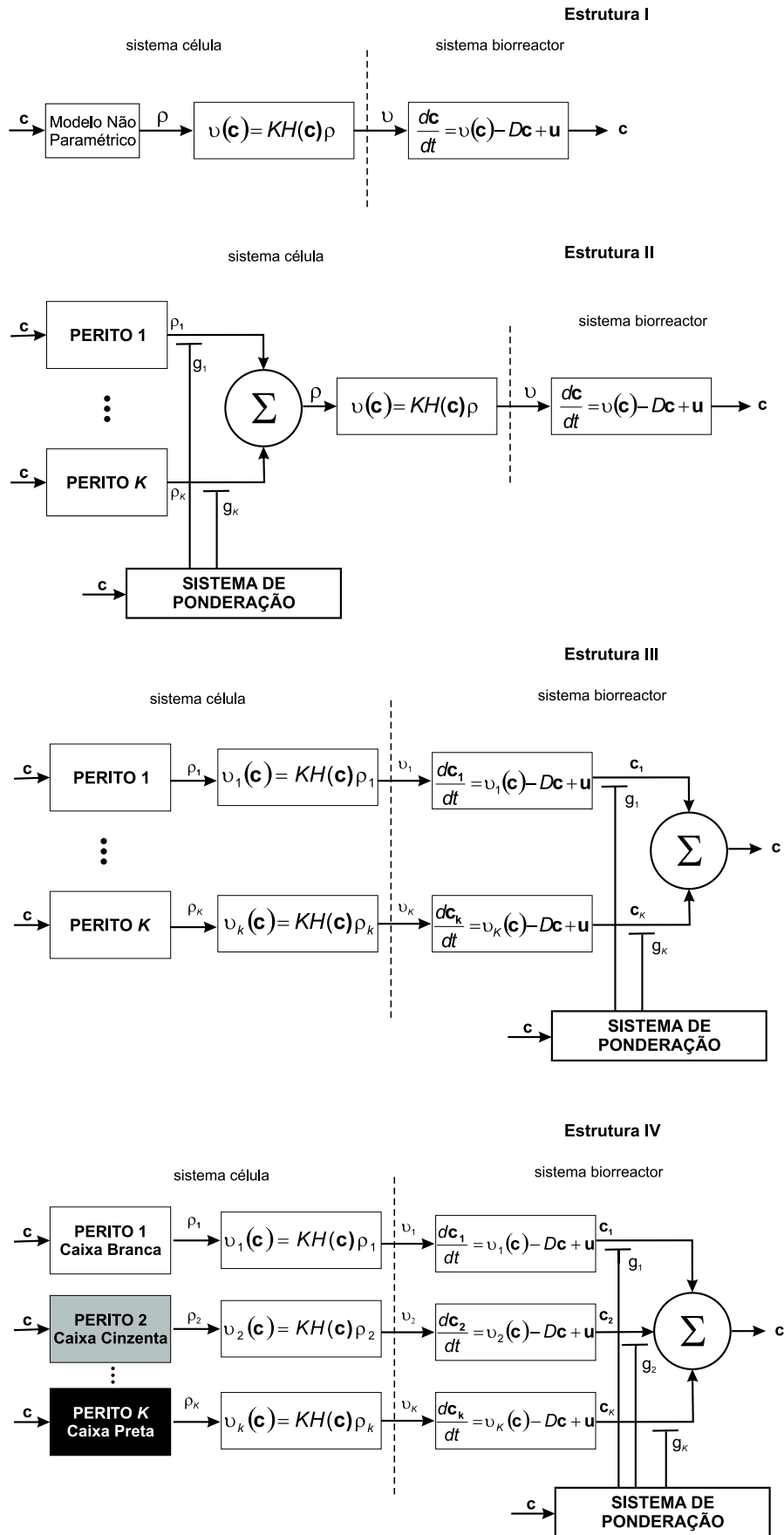


Figura 8.1: Estruturas híbridas estudadas neste trabalho.

3. desde que todas as funções do modelo que representa ρ sejam contínuas, diferenciáveis e limitadas, a análise de estabilidade entrada limitada saída limitada (BIBO) pode ser efectuada e, ainda mais importante, as sensibilidades dos parâmetros podem ser calculadas.

Relativamente a este tipo de estrutura concluiu-se que:

- o uso de métodos de integração robustos provou ser essencial para a identificação dos modelos não paramétricos;
- observou-se que a capacidade de descrever várias fermentações em simultâneo depende fortemente do grau de exactidão do método de integração empregue.

Discutiram-se as condições de estabilidade BIBO para a estrutura A e propôs-se uma nova configuração para o modelo híbrido que garante a positividade das concentrações em concordância com a realidade física.

A estrutura II representada na Figura 8.1 apresenta a vantagem de incorporar algum conhecimento a priori sobre processos biotecnológicos específicos onde ocorrem transições metabólicas. Nesta estrutura, o termo cinético desconhecido é modelizado por uma rede de mistura de peritos, onde cada perito descreve cada estado metabólico.

Primeiro estudou-se o sistema célula no Capítulo 5 e no Capítulo 6 desenvolveu-se o modelo completo onde a mistura dos peritos é feita ao nível das cinéticas.

A estrutura III da Figura 8.1 representa o modelo sistema célula mais o sistema biorreactor onde a mistura de peritos é feita ao nível das concentrações. Este tipo de estrutura foi estudada no Capítulo 6.

Os peritos das estruturas II e III são sempre do mesmo tipo que no caso deste trabalho são redes MLP.

Com o intuito de obter soluções mais flexíveis foi adoptado o sistema de ponderação Gaussiano para as estruturas II e III.

Com a introdução da rede de mistura de peritos para descrever o sistema célula chegaram-se às seguintes conclusões:

- este tipo de estrutura híbrida mimetiza o comportamento cinético de determinados processos biológicos;
- demonstrou-se que detecta a transição entre estados metabólicos;

- demonstrou-se que cada perito é capaz de descrever cada estado metabólico;
- do ponto de vista estrutural tem a vantagem adicional de ter uma correspondência com a realidade física.

Demonstrou-se, portanto, que com este tipo de estruturas suportadas por redes de mistura de peritos se consegue extrair mais informação acerca do processo a partir de dados, sejam eles simulados ou experimentais.

Da comparação entre as estruturas I e II da Figura 8.1 relativamente apenas ao sistema célula concluiu-se que:

- a rede mistura de peritos (ME) e a rede MLP supera sistematicamente a rede RBF em termos do rácio exactidão do modelo/número de parâmetros;
- o tempo de computação quando se utiliza uma rede de mistura de peritos é bastante superior dada a complexidade deste tipo de estruturas;
- a rede Mistura de Peritos (ME) se treinada com o algoritmo Esperança-Maximização (EM) é capaz de detectar a transição entre estados metabólicos sem falhar;
- a rede Mistura de Peritos (ME) tem a vantagem adicional de os peritos empregues desenvolverem capacidades individuais para descrever os estados metabólicos individualmente;
- a rede Mistura de Peritos (ME) é capaz de descrever com mais exactidão que a rede MLP as cinéticas na vizinhança das transições metabólicas.

A evolução natural deste tipo de estruturas culmina na estrutura IV da Figura 8.1 onde os peritos representam diferentes tipos de conhecimento com diferentes níveis de sofisticação. Este tipo de estrutura foi estudada no Capítulo 7 e do ponto de vista do desenho é semelhante à estrutura III mas difere essencialmente na definição dos peritos. A estrutura IV suporta diferentes tipos de peritos: do tipo caixa branca, do tipo caixa cinzenta ou do tipo caixa preta. Propôs-se um novo sistema de ponderação baseado em funções gaussianas para a estrutura IV.

As principais conclusões a que se chegaram sobre este último tipo de estrutura híbrida foram:

- introdução dum método de ponderação de modelos híbridos inovador independente do processo em estudo;

- este método tem a grande vantagem de entrar em consideração com a exactidão de cada modelo em diferentes regiões do espaço das entradas (problema que tinha ficado em aberto no Capítulo 2);
- garante-se que em cada instante o modelo mais exacto é usado para calcular a saída final;
- o uso combinado de peritos de diferentes tipos levou a melhoramentos na capacidade de predição comparado com os desempenhos individuais de cada perito;
- integração mais flexível de diferentes formas de conhecimento.

8.2 Identificação de parâmetros

Estudaram-se duas estratégias de identificação de parâmetros:

1. estratégia I: minimização directa no sentido dos mínimos quadrados dos erros entre as taxas de reacção medidas e modelizadas;
2. estratégia II: minimização directa dos erros nas concentrações no sentido dos mínimos quadrados empregando um algoritmo de SQP com os gradientes calculados pelas equações das sensibilidades.

Provou-se que a melhor estratégia de identificação é a estratégia II e também se demonstrou que pode ser aplicada com sucesso com dados experimentais esparsos e com ruído em oposição à estratégia I.

Estas estratégias de identificação de parâmetros foram aplicadas à estrutura I.

Relativamente às estruturas suportadas por redes de mistura de peritos o algoritmo de identificação adoptado está ligado intrinsecamente à natureza da estrutura da rede e das características da solução perspectivada. Como o objectivo é obter soluções competitivas adoptou-se a formulação da máxima verosimilhança com o algoritmo da Esperança-Maximização (EM).

Para a estrutura II aplicou-se o algoritmo EM standard onde a estratégia de identificação de parâmetros de cada perito foi a estratégia I. Para as estruturas III e IV aplicou-se a estratégia II na identificação dos parâmetros de cada perito.

Da comparação entre as estruturas II e III concluiu-se que:

- se obtém um modelo mais exacto, quando se aplica a estratégia de identificação baseada na minimização dos erros nas concentrações na identificação dos parâmetros de cada perito, isto é, com a estrutura III.

Note-se que todos os modelos híbridos foram validados. A técnica da validação cruzada foi a técnica adoptada para evitar o sobreajustamento do modelo aos dados, isto é, para evitar que o modelo se correlacionasse com o ruído do conjunto de dados usado para o treino.

8.3 Aplicabilidade aos processos biotecnológicos

Estes métodos foram aplicados aos seguintes processos biotecnológicos:

- a estrutura I foi aplicada ao processo de produção de proteína recombinante e ao processo de produção de fermento de padeiro ambos com dados simulados;
- a estrutura II (apenas a parte do sistema célula) foi aplicada ao processo de produção de fermento de padeiro com dados simulados e experimentais e ao processo de remoção de fósforo de águas residuais com dados simulados;
- a estrutura II (modelo completo) e a estrutura III foi aplicada ao processo de produção de Polihidroxialcanoatos (plástico biodegradável sintetizado por uma bactéria que tem propriedades similares ao polipropileno) com dados experimentais;
- a estrutura IV foi aplicada ao processo de produção de fermento de padeiro com dados experimentais.

Demonstrou-se que a técnica da modelização híbrida aplicada a processos biotecnológicos:

- tem a capacidade de descrever com exactidão o estado do processo;
- tem custos de desenvolvimento inferiores aos dos modelos mecanísticos;
- tem vantagem relativamente à técnicas puramente não paramétricas como já demonstrado também noutros trabalhos;
- é uma estratégia de modelização com elevada razão benefício/custo para bioprocessos.

8.4 Desenvolvimentos futuros

Esta tese deixa naturalmente em aberto algumas linhas de investigação, nomeadamente na demonstração da optimização dinâmica e controlo suportada por modelos híbridos.

Seria também importante construir uma aplicação computacional de fácil utilização para desenvolver modelos híbridos para optimização e controlo com taxas de benefício/custo elevadas.

Referências

- Bastin, G. e Dochain, D. (1990), *On-Line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam.
- Chen, L., Bernard, O., Bastin, G. e Angelov, P. (2000), 'Hybrid modelling of biotechnological processes using neural networks', *Control Engineering Practice*, **8**(7), 821–827.
- Dochain, D., Perrier, M. e Ydstie, B. E. (1992), 'Asymptotic observers for stirred tank reactors', *Chemical Engineering Science*, **47**(15-16), 4167–4177.
- Feyo de Azevedo, S., Dahm, B. e Oliveira, F. R. (1997), 'Hybrid modelling of biochemical processes: A comparison with the conventional approach', *Computers & Chemical Engineering*, **21**, S751–S756.
- Montague, G. e Morris, J. (1994), 'Neural-network contributions in biotechnology', *Trends in Biotechnology*, **12**(8), 312–324.
- Peres, J., Oliveira, R. e de Azevedo, S. F. (2001), 'Knowledge based modular networks for process modelling and control', *Computers & Chemical Engineering*, **25**(4-6), 783–791.
- Preusting, H., Noordover, J., Simutis, R. e Lubbert, A. (1996), 'The use of hybrid modelling for the optimization of the penicillin fermentation process', *Chimia*, **50**(9), 416–417.
- Psichogios, D. C. e Ungar, L. H. (1992), 'A hybrid neural network-1st principles approach to process modeling', *AIChE Journal*, **38**(10), 1499–1511.
- Schubert, J., Simutis, R., Dors, M., Havlik, I. e Lubbert, A. (1994a), 'Bioprocess optimization and control application of hybrid modeling', *Journal of Biotechnology*, **35**(1), 51–68.

- Schubert, J., Simutis, R., Dors, M., Havlik, I. e Lubbert, A. (1994b), 'Hybrid modeling of yeast production processescombination of a-priori knowledge on different levels of sophistication', *Chemical Engineering & Technology*, **17**(1), 10–20.
- Simutis, R., Oliveira, R., Manikowski, M., de Azevedo, S. F. e Lubbert, A. (1997), 'How to increase the performance of models for process optimization and control', *Journal of Biotechnology*, **59**(1-2), 73–89.
- Thompson, M. L. e Kramer, M. A. (1994), 'Modeling chemical processes using prior knowledge and neural networks', *AIChE Journal*, **40**(8), 1328–1340.
- van Can, H. J. L., Braake, H. A. B. T., Hellinga, C., Luyben, K. C. A. M. e Heijnen, J. J. (1999), 'An efficient model development strategy for bioprocesses based on neural networks in macroscopic balances: Part ii', *Biotechnology and Bioengineering*, **62**(6), 666–680.
- van Can, H. J. L., teBraake, H. A. B., Dubbelman, S., Hellinga, C., Luyben, K. C. A. M. e Heijnen, J. J. (1998), 'Understanding and applying the extrapolation properties of serial gray-box models', *AIChE Journal*, **44**(5), 1071–1089.

Lista de referências utilizadas na Tese

Nesta tese adotou-se o princípio de referir no fim de cada capítulo os trabalhos neles citados. Para possível facilidade de apreciação lista-se neste anexo todas as referências citadas.

Anderson, J. S., McAvoy, T. J. e Hao, O. J. (2000), 'Use of hybrid models in wastewater systems', *Industrial & Engineering Chemistry Research*, **39**(6), 1694–1704.

Babuska, R., Verbruggen, H. B. e van Can, H. J. L. (1999), 'Fuzzy modeling of enzymatic penicillin-g conversion', *Engineering Applications Of Artificial Intelligence*, **12**(1), 79–92.

Barlow, T. W. (1995), 'Feedforward neural networks for secondary structure prediction', *Journal of Molecular Graphics*, **13**(3), 175–183.

Bastin, G. e Dochain, D. (1990), *On-Line Estimation and Adaptive Control of Bioreactors*, Elsevier, Amsterdam.

Beccari, M., Majone, M., Massanisso, P. e Ramadori, R. (1998), 'A bulking sludge with high storage response selected under intermittent feeding', *Water Research*, **32**(11), 3403–3413.

Beun, J. J., Dircks, K., Van Loosdrecht, M. C. M. e Heijnen, J. J. (2002), 'Poly-*beta*-hydroxybutyrate metabolism in dynamically fed mixed microbial cultures', *Water Research*, **36**(5), 1167–1180.

Beun, J. J., Paletta, F., Van Loosdrecht, M. C. M. e Heijnen, J. J. (2000), 'Stoichiometry and kinetics of poly-*beta*-hydroxybutyrate metabolism in aerobic, slow growing, activated sludge cultures', *Biotechnology and Bioengineering*, **67**(4), 379–389.

Bhat, N. e Mcavoy, T. J. (1990), 'Use of neural nets for dynamic modeling and control of chemical process systems', *Computers & Chemical Engineering*, **14**(4-5), 573–583.

- Bishop, C. M. (1995), *Neural Networks for Pattern Recognition*, Oxford University Press.
- Braake, H. A. B. T., van Can, H. J. L. e Verbruggen, H. B. (1998), 'Semi-mechanistic modeling of chemical processes with neural networks', *Engineering Applications Of Artificial Intelligence*, **11**(4), 507–515.
- Breiman, L., Friedman, J. H., Olshen, R. A. e Stone, C. J. (1984), *Classification and Regression Trees*, Belmont, CA: Wadsworth.
- Cardoso, M. F., Salcedo, R. L. e Deazevedo, S. F. (1994), 'Nonequilibrium simulated annealing a faster approach to combinatorial minimization', *Industrial & Engineering Chemistry Research*, **33**(8), 1908–1918.
- Carucci, A., Dionisi, D., Majone, M., Rolle, E. e Smurra, P. (2001), 'Aerobic storage by activated sludge on real wastewater', *Water Research*, **35**(16), 3833–3844.
- Chen, L., Bernard, O., Bastin, G. e Angelov, P. (2000), 'Hybrid modelling of biotechnological processes using neural networks', *Control Engineering Practice*, **8**(7), 821–827.
- Chinrungrueng, C. (1993), *Evaluation of heterogenous Architectures for Artificial Neural Networks*, Tese de Doutorado, University of California at Berkeley.
- Chinrungrueng, C. e Sequin, C. H. (1995), 'Optimal adaptive k-means algorithm with dynamic adjustment of learning rate', *IEEE Transactions On Neural Networks*, **6**(1), 157–169.
- Cybenko, G. (1989), 'Approximation by superpositions of a sigmoidal function', *Mathematics of Control, Signals, and Systems*, **2**, 303–314.
- Daigger, G. T. e Grady, C. P. L. (1982), 'An assessment of the role of physiological adaptation in the transient-response of bacterial cultures', *Biotechnology and Bioengineering*, **24**(6), 1427–1444.
- Dailey, M. N. e Cottrell, G. W. (1999), 'Organization of face and object recognition in modular neural network models', *Neural Networks*, **12**(7-8), 1053–1073.
- Dempster, A. P., Laird, N. M. e Rubin, D. B. (1977), 'Maximum likelihood from incomplete data via em algorithm', *Journal of The Royal Statistical Society Series B-Methodological*, **39**(1), 1–38.
- Dochain, D., Perrier, M. e Ydstie, B. E. (1992), 'Asymptotic observers for stirred tank reactors', *Chemical Engineering Science*, **47**(15-16), 4167–4177.

- Edgar, T. F. (1996), 'Modelling and control - back to the future, part i', *CAST Communications*, **19**(1), 7–12.
- Eikens, B. e Karim, M. N. (1999), 'Process identification with multiple neural network models', *International Journal Of Control*, **72**(7-8), 576–590.
- Ferreira, E. (1995), *Identificação e Controlo Adaptativo de Processos Biotecnológicos*, Tese de Doutoramento, Faculdade de Engenharia da Universidade do Porto.
- Feyo de Azevedo, S., Dahm, B. e Oliveira, F. R. (1997), 'Hybrid modelling of biochemical processes: A comparison with the conventional approach', *Computers & Chemical Engineering*, **21**, S751–S756.
- Feyo de Azevedo, S., Oliveira, R. e Sonnleitner, B. (2001), *Novel Multiphase Bioreactors*, Harwood Academic Publishers, UK, Capítulo 3: New Metodologies for Multiphase Bioreactors: Data Acquisition, Modelling and Control.
- Georgieva, P., Meireles, M. J. e de Azevedo, S. F. (2003), 'Knowledge-based hybrid modelling of a batch crystallisation when accounting for nucleation, growth and agglomeration phenomena', *Chemical Engineering Science*, **58**(16), 3699–3713.
- Glasse, J., Ignova, M., Ward, A. C., Montague, G. A. e Morris, A. J. (1997), 'Bioprocess supervision: Neural networks and knowledge based systems', *Journal of Biotechnology*, **52**(3), 201–205.
- Gujer, W., Henze, M., Mino, T. e van Loosdrecht, M. (1999), 'Activated sludge model no. 3', *Water Science and Technology*, **39**(1), 183–193.
- Haykin, S. (1994), *Neural Networks: A comprehensive foundation*, Macmillan College Publishing Company, Inc.
- Haykin, S. (1999), *Neural Networks: A comprehensive foundation*, Prentice Hall, Inc., 2 Edição.
- Henze, M., Gujer, W., Mino, T., Matsuo, T., Wentzel, M. C., Marais, G. V. R. e Van Loosdrecht, M. C. M. (1999), 'Activated sludge model no.2d, asm2d', *Water Science and Technology*, **39**(1), 165–182.
- Henze, M., Gujer, W., Mino, T. e van Loosdrecht, M. E. (Editores) (2000), *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3; Scientific and Technical Report 9*, IWA Publishing, London.

- Hinton, G., Revow, M. e Dayan, P. (1995), 'Recognizing handwritten digits using mixture of linear models', em G. Tesauro, D. Touretzky e T. Leen (Editores), *Advances in Neural Information Processing Systems*, The MIT Press, volume 7.
- Hitzmann, B., Lubbert, A. e Schugerl, K. (1992), 'An expert system approach for the control of a bioprocess .1. knowledge representation and processing', *Biotechnology and Bioengineering*, **39**(1), 33–43.
- Horiuchi, J. e Hiraga, K. (1999), 'Industrial application of fuzzy control to large-scale recombinant vitamin b-2 production', *Journal Of Bioscience And Bioengineering*, **87**(3), 365–371.
- Hornik, K., Stinchcombe, M. e White, H. (1989), 'Multilayer feedforward networks are universal approximators', *Neural Networks*, **2**(5), 359–366.
- Hu, Y. H., Palreddy, S. e Tompkins, W. J. (1997), 'A patient-adaptable ecg beat classifier using a mixture of experts approach', *IEEE Transactions on Biomedical Engineering*, **44**(9), 891–900.
- Hunt, K. J., Sbarbaro, D., Zbikowski, R. e Gawthrop, P. J. (1992), 'Neural networks for control-systems: a survey', *Automatica*, **28**(6), 1083–1112.
- Jacobs, R. A. e Jordan, M. I. (1991), 'A competitive modular connectionist architecture', em J. M. R.P. Lippman e D. Touretzky (Editores), *Advances in Neural Information Processing Systems*, CA Morgan Kaufmann, San Mateo, volume 3, Páginas 767–773.
- Jacobs, R. A. e Jordan, M. I. (1993), 'Learning piecewise control strategies in a modular neural-network architecture', *IEEE Transactions on Systems Man and Cybernetics*, **23**(2), 337–345.
- Jacobs, R. A., Jordan, M. I. e Barto, A. G. (1991a), 'Task decomposition through competition in a modular connectionist architecturethe what and where vision tasks', *Cognitive Science*, **15**(2), 219–250.
- Jacobs, R. A., Jordan, M. I., Nowlan, S. J. e Hinton, G. E. (1991b), 'Adaptive mixtures of local experts', *Neural Computation*, **3**, 79–87.
- Johansen, T. A. e Foss, B. A. (1997), 'Operating regime based process modeling and identification', *Computers & Chemical Engineering*, **21**(2), 159–176.
- Jordan, M. I. e Jacobs, R. A. (1994), 'Hierarchical mixtures of experts and the em algorithm', *Neural Computation*, **6**(2), 181–214.

- Jordan, M. I. e Xu, L. (1995), 'Convergence results for the em approach to mixtures of experts architectures', *Neural Networks*, **8**(9), 1409–1431.
- Karama, A., Bernard, O., Genovesi, A., Dochain, D., Benhammou, A. e Steyer, J. P. (2001a), 'Hybrid modelling of anaerobic wastewater treatment processes', *Water Science and Technology*, **43**(1), 43–50.
- Karama, A., Bernard, O., Gouze, J. L., Benhammou, A. e Dochain, D. (2001b), 'Hybrid neural modelling of an anaerobic digester with respect to biological constraints', *Water Science and Technology*, **43**(7), 1–8.
- Kasabov, N. K. (1996), *Foundations of neural Networks, Fuzzy Systems, and Knowledge Engineering*, Bradford Books, The MIT Press.
- Kohonen, T. (1995), *Self-Organizing Maps*, Springer, Heidelberg.
- Kosko, B. (1992), *Neural networks and fuzzy systems: a dynamical system approach to machine intelligence*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Krishna, C. e Van Loosdrecht, M. C. M. (1999a), 'Effect of temperature on storage polymers and settleability of activated sludge', *Water Research*, **33**(10), 2374–2382.
- Krishna, C. e Van Loosdrecht, M. C. M. (1999b), 'Substrate flux into storage and growth in relation to activated sludge modeling', *Water Research*, **33**(14), 3149–3161.
- Lauret, P., Boyer, H. e Gatina, J. (2000), 'Hybrid modelling of a sugar boiling process', *Control Engineering Practice*, **8**, 299–310.
- Lee, D. S., Vanrolleghem, P. A. e Park, J. M. (2005), 'Parallel hybrid modeling methods for a full-scale cokes wastewater treatment plant', *Journal Of Biotechnology*, **115**(3), 317–328.
- Lee, S. L., Jeon, C. O., Park, J. M. e Chang, K. S. (2002), 'Hybrid neural network modeling of a full-scale industrial wastewater treatment process', *Biotechnology and Bioengineering*, **78**(6), 670–682.
- Leonard, J. e Kramer, M. A. (1990), 'Improvement of the backpropagation algorithm for training neural networks', *Computers & Chemical Engineering*, **14**(3), 337–341.
- Leonard, J. A., Kramer, M. A. e Ungar, L. H. (1992), 'A neural network architecture that computes its own reliability', *Computers & Chemical Engineering*, **16**(9), 819–835.
- Ljung, J. (1987), *System Identification - Theory for the User*, Prentice-Hall.

- Lodish, H., Berk, A., Matsudaira, P., Kaiser, C. A., Krieger, M., Scott, M. P., Zipursky, L. e Darnell, J. (2000), *Molecular Cell Biology*, W.H. Freeman.
- Lubbert, A. e Simutis, R. (1994), 'Using measurement data in bioprocess modeling and control', *Trends in Biotechnology*, **12**(8), 304–311.
- McLachlan, G. H. e Basford, K. E. (1988), *Mixture Models: Inference and Application to Clustering*, New York: Marcel Dekker.
- Melin, P., Felix, C. e Castillo, O. (2005), 'Face recognition using modular neural networks and the fuzzy sugeno integral for response integration', *International Journal of Intelligent Systems*, **20**(2), 275–291.
- Miller, D., Rao, A. V., Rose, K. e Gersho, A. (1996), 'A global optimization technique for statistical classifier design', *IEEE Transactions on Signal Processing*, **44**(12), 3108–3122.
- Moerlan, P. (2000), *Mixture Models for Unsupervised and Supervised Learning*, Tese de Doutorado, Computer Science Department, Swiss Federal Institute of Technology at Lausanne (EPFL).
- Molga, E. J. (2003), 'Neural network approach to support modelling of chemical reactors: problems, resolutions, criteria of application', *Chemical Engineering And Processing*, **42**(8-9), 675–695.
- Moller, M. F. (1993), 'A scaled conjugate-gradient algorithm for fast supervised learning', *Neural networks*, **6**(4), 525–533.
- Montague, G. e Morris, J. (1994), 'Neural-network contributions in biotechnology', *Trends in Biotechnology*, **12**(8), 312–324.
- Moody, J. e Darken, C. J. (1989), 'Fast learning in networks of locally-tuned processing units', *Neural Computation*, **1**, 281 – 294.
- Narendra, K. e Parthasarathy, K. (1990), 'Identification and control of dynamical systems using neuralnetworks', *IEEE Transactions on Neural Networks*, **1**(1), 4–27.
- Oliveira, F. M. (1997), *Monitorização e Controlo de Fermentadores: Aplicação ao Fermento de Padeiro*, Tese de Doutorado, Faculdade de Engenharia da Universidade do Porto.
- Oliveira, R. (1998), *Supervision, Control and Optimization of Biotechnological Processes Based on Hybrid Models*, Tese de Doutorado, Martin-Luther-Universitat Halle-Wittenberg.

- Oliveira, R., Ferreira, E. C. e de Azevedo, S. F. (2002), 'Stability, dynamics of convergence and tuning of observer-based kinetics estimators', *Journal of Process Control*, **12**(2), 311–323.
- Oliveira, R., Peres, J. e Feye de Azevedo, S. (2000), 'Efficient knowledge integration methods for improved bioreactor operation', em *4th Portuguese Conference on Automatic Control (Controlo'2000)*, Guimarães, Portugal, Páginas 214–218.
- Oliveira, R., Peres, J. e Feye de Azevedo, S. (2005), 'Hybrid modelling of fermentation processes using artificial neural networks: A study on identification and stability', em M. Pons e J. F. M. van Impe (Editores), *Computer Applications in Biotechnology 2004*, Elsevier (ISBN: 0-08-044251-X), Páginas 195 – 200.
- Park, S. e Ramirez, W. F. (1988), 'Optimal production of secreted protein in fed-batch reactors', *AIChE Journal*, **34**(9), 1550–1558.
- Park, S. e Ramirez, W. F. (1989), 'Dynamics of foreign protein secretion from *Saccharomyces-cerevisiae*', *Biotechnology and Bioengineering*, **33**(3), 272–281.
- Peng, F. C., Jacobs, R. A. e Tanner, M. A. (1996), 'Bayesian inference in mixtures-of-experts and hierarchical mixtures-of-experts models with an application to speech recognition', *Journal of the American Statistical Association*, **91**(435), 953–960.
- Peres, J., Oliveira, R. e de Azevedo, S. F. (2001), 'Knowledge based modular networks for process modelling and control', *Computers & Chemical Engineering*, **25**(4-6), 783–791.
- Peres, J., Oliveira, R. e de Azevedo, S. F. (2003), 'Modelling cells reaction kinetics with artificial neural networks: A comparison of three network architectures', em A. Kraslawski e I. Turunen (Editores), *European Symposium On Computer Aided Process Engineering - 13*, Elsevier Science Bv, volume 14 de *Computer-Aided Chemical Engineering*, Páginas 839–844.
- Peres, J., Oliveira, R. e Feye de Azevedo, S. (2005a), 'Hybrid modelling of fermentation processes: A study on the use of modular neural networks for modelling cells reaction kinetics', em M. Pons e J. F. M. van Impe (Editores), *Computer Applications in Biotechnology 2004*, Elsevier (ISBN: 0-08-044251-X), Páginas 293 – 298.
- Peres, J., Oliveira, R. e Feye de Azevedo, S. (2005b), 'A study on the application of modular neural networks for modelling cell reaction kinetics', em *preparação para publicação*.

- Peres, J., Oliveira, R., Serafim, L. S., Lemos, P., Reis, M. A. e de Azevedo, S. F. (2004), 'Hybrid modelling of a pha production process using modular neural networks', em A. Barbosa-Póvoa e H. Matos (Editores), *European Symposium On Computer-Aided Process Engineering - 14*, Elsevier Science Bv, volume 18 de *Computer-Aided Chemical Engineering*, Páginas 733–738.
- Poggio, T. e Girosi, F. (1990), 'Networks for approximation and learning', *Proceedings of the IEEE*, **78**(9), 1481–1497.
- Pollard, J. F., Broussard, M. R., Garrison, D. B. e San, K. Y. (1992), 'Process identification using neural networks', *Computers & Chemical Engineering*, **16**(4), 253–270.
- Pomerleau, Y. e Perrier, M. (1990), 'Estimation of multiple specific growth-rates in bioprocesses', *AIChE Journal*, **36**(2), 207–215.
- Preusting, H., Noordover, J., Simutis, R. e Lubbert, A. (1996), 'The use of hybrid modelling for the optimization of the penicillin fermentation process', *Chimia*, **50**(9), 416–417.
- Psichogios, D. C. e Ungar, L. H. (1992), 'A hybrid neural network-1st principles approach to process modeling', *AIChE Journal*, **38**(10), 1499–1511.
- Qin, S. Z., Su, H. T. e Mcavoy, T. J. (1992), 'Comparison of 4 neural net learning-methods for dynamic system-identification', *IEEE Transactions on Neural Networks*, **3**(1), 122–130.
- Ramamurti, V. e Ghosh, J. (1999), 'Structurally adaptive modular networks for nonstationary environments', *IEEE Transactions on Neural Networks*, **10**(1), 152–160.
- Rao, A. V., Miller, D., Rose, K. e Gersho, A. (1997), 'Mixture of experts regression modeling by deterministic annealing', *IEEE Transactions on Signal Processing*, **45**(11), 2811–2820.
- Reed, G. e Peppler, H. (1973), *Yeast Technology*, The AVI Publishing Company, Inc., Connecticut.
- Reiling, H. E., Laurila, H. e Fiechter, A. (1985), 'Mass-culture of escherichia-colimedium development for low and high-density cultivation of escherichia coli-b/r in minimal and complex media', *Journal of Biotechnology*, **2**(3-4), 191–206.
- Roubos, J. A., Krabben, P., Setness, M., Babuska, R., Heijnen, J. e Verbrugen, H. B. (1999), 'Hybrid model development for fed-batch bioprocesses combining physical

- equations with the metabolic network and black-box kinetics', em *6th Workshop on fuzzy systems*, Brunel University, Uxbridge, Páginas 231–239.
- Royce, P. N. (1993), 'A discussion of recent developments in fermentation monitoring and control from a practical perspective', *Critical Reviews in Biotechnology*, **13**(2), 117–149.
- Rumelhart, D. E., Hinton, G. E. e Williams, R. J. (1986), 'Learning internal representations by error propagation', em D. E. Rumelhart, J. L. McClelland e the PDP Research Group (Editores), *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Cambridge, MA: MIT Press, volume 1: Foundations, Páginas 318 – 362.
- Russell, N. T. e Bakker, H. H. C. (1997), 'Modular modelling of an evaporator for long-range prediction', *Artificial Intelligence in Engineering*, **11**(4), 347–355.
- Schubert, J., Simutis, R., Dors, M., Havlik, I. e Lubbert, A. (1994a), 'Bioprocess optimization and control application of hybrid modeling', *Journal of Biotechnology*, **35**(1), 51–68.
- Schubert, J., Simutis, R., Dors, M., Havlik, I. e Lubbert, A. (1994b), 'Hybrid modeling of yeast production processescombination of a-priori knowledge on different levels of sophistication', *Chemical Engineering & Technology*, **17**(1), 10–20.
- Scott, G. M. e Harmon Ray, W. (1993), 'Creating efficient nonlinear neural network process models that allow model interpretation', *Journal of Process Control*, **3**(3), 163–178.
- Senger, R. S. e Karim, M. N. (2003), 'Neural-network-based identification of tissue-type plasminogen activator protein production and glycosylation in cho cell culture under shear environment', *Biotechnology Progress*, **19**(6), 1828–1836.
- Serafim, L. S., Lemos, P. C., Oliveira, R. e Reis, M. A. M. (2004), 'Optimization of polyhydroxybutyrate production by mixed cultures submitted to aerobic dynamic feeding conditions', *Biotechnology and Bioengineering*, **87**(2), 145–160.
- Simutis, R., Havlik, I. e Lubbert, A. (1993), 'Fuzzy-aided neural network for real-time state estimation and process prediction in the alcohol formation step of production-scale beer brewing', *Journal of Biotechnology*, **27**(2), 203–215.
- Simutis, R., Havlik, I., Schneider, F., Dors, M. e Lübbert, A. (1995), 'Artificial neural networks of improved reliability for industrial process supervision', em *Preprints*

- of the 6th Int. Conference on Computer Applications in Biotechnology, Garmisch-Partenkirchen, Germany, Páginas 59–65.
- Simutis, R., Oliveira, R., Manikowski, M., de Azevedo, S. F. e Lubbert, A. (1997), 'How to increase the performance of models for process optimization and control', *Journal of Biotechnology*, **59**(1-2), 73–89.
- Sjoberg, J., Zhang, Q. H., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P. Y., Hjalmarsson, H. e Juditsky, A. (1995), 'Nonlinear black-box modeling in system identification: A unified overview', *Automatica*, **31**(12), 1691–1724.
- Sonnleitner, B. (1999), 'Bioanalysis and biosensors for bioprocess monitoring', *Advances in Biochemical Engineering/Biotechnology* 66, (volume editor).
- Sonnleitner, B. e Kappeli, O. (1986), 'Growth of *Saccharomyces-cerevisiae* is controlled by its limited respiratory capacity formulation and verification of a hypothesis', *Biotechnology and Bioengineering*, **28**(6), 927–937.
- Sterbacek, Z. e Votruba, J. (1993), 'An expert system applied to the control of an industrial-scale bioreactor', *Chemical Engineering Journal and The Biochemical Engineering Journal*, **51**(2), B35–B42.
- Sugeno, M. (1985), *Industrial applications of fuzzy control*, North-Holland, Amsterdam.
- Sá-Correia, I., Moreira, L. e Fialho, A. (2003), 'Engenharia genética', em N. Lima e M. Mota (Editores), *Biotecnologia: Fundamentos e Aplicações*, LIDEL - Edições Técnicas, Lda., Páginas 125 – 161.
- Third, K. A., Newland, M. e Cord-Ruwisch, R. (2003), 'The effect of dissolved oxygen on phb accumulation in activated sludge cultures', *Biotechnology and Bioengineering*, **82**(2), 238–250.
- Thompson, M. L. e Kramer, M. A. (1994), 'Modeling chemical processes using prior knowledge and neural networks', *AIChE Journal*, **40**(8), 1328–1340.
- Titterton, D. M., Smith, A. F. M. e Makov, U. E. (1985), *Analysis of Finite Mixture Distributions*, New York: Wiley.
- van Aalst-van Leeuwen, M. A., Pot, M. A., van Loosdrecht, M. C. M. e Heijnen, J. J. (1997), 'Kinetic modeling of poly(beta-hydroxybutyrate) production and consumption by *Paracoccus pantotrophus* under dynamic substrate supply', *Biotechnology and Bioengineering*, **55**(5), 773–782.

- van Can, H. J. L., Braake, H. A. B. T., Hellinga, C., Luyben, K. C. A. M. e Heijnen, J. J. (1999), 'An efficient model development strategy for bioprocesses based on neural networks in macroscopic balances: Part ii', *Biotechnology and Bioengineering*, **62**(6), 666–680.
- van Can, H. J. L., teBraake, H. A. B., Dubbelman, S., Hellinga, C., Luyben, K. C. A. M. e Heijnen, J. J. (1998), 'Understanding and applying the extrapolation properties of serial gray-box models', *AIChE Journal*, **44**(5), 1071–1089.
- van Loosdrecht, M. C. M. e Heijnen, J. J. (2002), 'Modelling of activated sludge processes with structured biomass', *Water Science and Technology*, **45**(6), 13–23.
- Villermaux, J. (1996), 'Future prospects for chemical engineering research and technology', *Chem. Tech. Europe*, 21–23.
- Walsh, G. (2002), *Proteins: Biochemistry and Biotechnology*, John Wiley Sons.
- Wang, L. X. (1994), *Adaptive fuzzy systems and control: design and stability analysis*, Prentice-Hall, Englewood Cliffs, New Jersey.
- Waterhouse, S. R. (1993), *Speech recognition using hierarchical mixture of experts*, Tese de Mestrado, Cambridge University Engineering Department, Trumpington Street, Cambridge CB2 1PZ, UK.
- Weigend, A. S., Mangeas, M. e Srivastava, A. N. (1995), 'Nonlinear gated experts for time series: Discovering regimes and avoiding overfitting', *International Journal of Neural Systems*, **6**(4), 373–399.
- Xu, L. e Jordan, M. I. (1996), 'On convergence properties of the em algorithm for gaussian mixtures', *Neural Computation*, **8**(1), 129–151.
- Xu, L., Jordan, M. I. e Hinton, G. E. (1995), 'An alternative model for mixture of experts', em G. Tesauro, D. S. Touretzky e T. K. Leen (Editores), *Advances in Neural Information Processing Systems*, MIT Press, volume 7, Páginas 633–640.
- Ye, K., Fujioka, K. e Shimizu, K. (1994), 'Efficient control of fed-batch baker's yeast cultivation based on neural network', *Process Control and Quality*, **5**(4), 245–250.
- Zhao, H., Hao, O. J., McAvoy, T. J. e Chang, C. (1997), 'Modeling nutrient dynamics in sequencing batch reactor', *Journal of Environmental Engineering*, **123**, 311–319.
- Zorzetto, L. F. M., Maciel, R. e Wolf-Maciel, M. R. (2000), 'Process modelling development through artificial neural networks and hybrid models', *Computers & Chemical Engineering*, **24**(2-7), 1355–1360.